N° 11F0027MIF au catalogue — N° 047
ISSN 1703-0412
ISBN 978-0-662-07537-0

Document de recherche

Série de documents de recherche sur l'analyse économique (AE)

Estimation de la PTF en présence de points aberrants et de points leviers : examen de l'ensemble de données KLEMS

par Ryan Macdonald

Division de l'analyse microéconomique
Immeuble R.-H.-Coats, 18-F 100, promenade Tunney's Pasture
Ottawa (Ontario) K1A 0T6

Téléphone: 1-800-263-1136

Statistique Canada Statistics Canada

Canada

Estimation de la PTF en présence de points aberrants et de points leviers : examen de l'ensemble de données KLEMS

par Ryan Macdonald

11F0027MIF n° 047
ISSN 1703-0412
ISBN 978-0-662-07537-0

Statistique Canada, Division de l'analyse microéconomique
Immeuble R.-H.-Coats, 18-F 100, promenade Tunney's Pasture
Ottawa (Ontario) K1A 0T6

Comment obtenir d'autres renseignements :
Service national de renseignements: 1-800-263-1136
Renseignements par courriel : infostats@statcan.ca

Décembre 2007

L'auteur désire remercier monsieur John Baldwin pour sa contribution constructive lors de la rédaction de ce document. Tout long du processus, l'auteur a bénéficié de ses conseils et de son aide. Les commentaires de Wulong Gu ainsi que les discussions avec celui-ci ont été grandement appréciés.

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2007

Tous droits réservés. Le contenu de la présente publication électronique peut être reproduit en tout ou en partie, et par quelque moyen que ce soit, sans autre permission de Statistique Canada, sous réserve que la reproduction soit effectuée uniquement à des fins d'étude privée, de recherche, de critique, de compte rendu ou en vue d'en préparer un résumé destiné aux journaux et/ou à des fins non commerciales. Statistique Canada doit être cité comme suit : Source (ou « Adapté de », s'il y a lieu) : Statistique Canada, année de publication, nom du produit, numéro au catalogue, volume et numéro, période de référence et page(s). Autrement, il est interdit de reproduire le contenu de la présente publication, ou de l'emmagasiner dans un système d'extraction, ou de le transmettre sous quelque forme ou par quelque moyen que ce soit, reproduction électronique, mécanique, photographique, pour quelque fin que ce soit, sans l'autorisation écrite préalable des Services d'octroi de licences, Division des services à la clientèle, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6.

This publication is available in English (Catalogue no. 11F0027MIE, no. 047).

Note de reconnaissance :
Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population, les entreprises, les administrations canadiennes et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques précises et actuelles.

Normes de service à la clientèle
Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle qui sont observées par les employés lorsqu'ils offrent des services à la clientèle. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.ca sous À propos de nous > Offrir des services aux Canadiens.

Série de documents de recherche sur l'analyse économique

La série de documents de recherche sur l'analyse économique permet de faire connaître les travaux de recherche effectués par le personnel du Secteur des études analytiques et des comptes nationaux, les boursiers invités et les universitaires associés. La série de documents de recherche a pour but de favoriser la discussion sur un éventail de sujets tels que les répercussions de la nouvelle économie, les questions de productivité, la rentabilité des entreprises, l'utilisation de la technologie, l'incidence du financement sur la croissance des entreprises, les fonctions de dépréciation, l'utilisation de comptes satellites, les taux d'épargne, le crédit-bail, la dynamique des entreprises, les estimations hédoniques, les tendances en matière de diversification et en matière d'investissements, les différences liées au rendement des petites et des grandes entreprises ou des entreprises nationales et multinationales ainsi que les estimations relatives à la parité du pouvoir d'achat. Les lecteurs de la série sont encouragés à communiquer avec les auteurs pour leur faire part de leurs commentaires, critiques et suggestions.

Les documents sont diffusés principalement au moyen d'Internet. Ils peuvent être téléchargés gratuitement sur Internet, à www.statcan.ca.

Tous les documents de recherche de la Série d'analyse économique font l'objet d'un processus de révision institutionnelle et d'évaluation par les pairs afin de s'assurer de leur conformité au mandat confié par le gouvernement à Statistique Canada en tant qu'agence statistique et de leur pleine adhésion à des normes de bonne pratique professionnelle, partagées par la majorité.

Les documents de cette série comprennent souvent des résultats provenant d'analyses statistiques multivariées ou d'autres techniques statistiques. Il faut noter que les conclusions de ces analyses sont sujettes à des incertitudes dans les estimations énoncées.

Le niveau d'incertitude dépendra de plusieurs facteurs : de la nature de la forme fonctionnelle de l'analyse multivariée utilisée; de la technique économétrique employée; de la pertinence des hypothèses statistiques sous-jacentes au modèle ou à la technique; de la représentativité des variables prises en compte dans l'analyse; et de la précision des données employées. Le processus de la revue des pairs vise à garantir que les documents dans les séries correspondent aux normes établies afin de minimiser les problèmes dans chacun de ces domaines.

Comité de révision des publications
Direction des études analytiques, Statistique Canada
18e étage, Immeuble R.-H.-Coats
Ottawa, Ontario K1A 0T6

Table des matières

Résumé **5**

Sommaire exécutif **6**

1. Introduction **8**

2. Données KLEMS **10**

3. Spécification du modèle **11**

4. Tests de racine unitaire **12**

4.1 Observations atypiques et tests de racine unitaire 13
4.2 Tests de racine unitaire des industries échantillonnées 14
4.3 Examen détaillé de l'industrie de la fabrication de textiles et de produits textiles 18
4.4 Tests de racine unitaire sur données de panel 21
4.5 Transformation des données 22

5. Spécification économétrique **25**

6. Estimation en présence d'observations atypiques **27**

6.1 Estimateur MCO 27
6.2 Estimateur S 28
6.3 Repérage des points aberrants et des points leviers 29

7. Estimations paramétriques de la PTF **30**

7.1 Estimations de la PTF pour les industries échantillonnées 30
7.2 Estimations de la PTF pour le panel d'industries 33

8. Conclusion **45**

Annexe A Points leviers et points aberrants selon l'équation de Cobb-Douglas **47**

Annexe B Points leviers et points aberrants selon l'équation de productivité du travail ... **52**

Bibliographie **57**

Résumé

Le présent document porte sur l'effet des observations atypiques présentes dans la base de données KLEMS (capital, travail, énergie, matières et services) et propose une méthode pour le contrer. Le niveau de désagrégation, la construction des données et les chocs économiques peuvent se traduire par des observations atypiques susceptibles d'influer sur les estimations et sur l'inférence si des précautions ne sont pas prises. Les prétests couramment utilisés, tels que les tests DFA (Dickey-Fuller amélioré) et KPSS (Kwaitkowski, Phillips, Schmidt et Shin), doivent être appliqués avec prudence dans ce contexte parce qu'ils sont sensibles aux points de données inhabituels. En outre, les méthodes les plus connues qui sont employées pour établir des estimations statistiques, comme la méthode des moindres carrés ordinaires, peuvent ne pas fonctionner correctement en présence d'observations atypiques. Afin de pallier ce problème, le document illustre une méthode robuste d'estimation des relations statistiques.

Mots-clés : productivité, points aberrants, série chronologique

Sommaire exécutif

La prudence est de mise au moment d'estimer des relations économiques à partir de la base de données KLEMS (capital, travail, énergie, matières et services). Bien que cet ensemble de données intègre des estimations de qualité de variables économiques construites à partir des tableaux des ressources et des emplois des Comptes nationaux, la présence d'observations atypiques, telles que des points aberrants et des points leviers, en constitue un aspect important.

Des observations atypiques peuvent être présentes pour plusieurs raisons. Premièrement, des changements apportés à la classification et à la méthodologie peuvent entraîner des ruptures dans le temps. Deuxièmement, la désagrégation des industries en classifications plus fines produit des données dont la cohérence peut être moindre. Troisièmement, les chocs macroéconomiques ou autres peuvent causer des mouvements abrupts et inhabituels dans les sources de données sous-jacentes.

Le présent document constitue une première étape dans le processus de traitement des observations atypiques. Il se divise en deux parties, la première présentant un examen des prétests de racine unitaire ou de stationnarité appliqués aux données KLEMS. Les résultats obtenus laissent entendre qu'en raison de la complexité des données et de la présence d'observations atypiques, les tests courants, qu'ils soient appliqués individuellement ou sur le panel d'industries, ne permettent peut-être pas une inférence adéquate.

La seconde partie porte sur la façon dont les observations atypiques influent sur les estimations paramétriques de la croissance de la productivité totale des facteurs (PTF) et sur la façon de les repérer. Lorsque des observations atypiques sont présentes, les méthodes d'estimation courantes, telles que la méthode des moindres carrés ordinaires (MCO), peuvent être moins efficaces. Ce document examine leur incidence en employant une technique d'estimation qui est moins sensible aux observations inhabituelles. Les résultats des deux méthodes d'estimation sont comparées afin d'illustrer l'incidence des points de données inhabituels.

La seconde partie du document fournit en outre des réponses à un certain nombre de questions concernant le traitement des observations atypiques :

- Quelles techniques peut-on utiliser pour traiter les observations atypiques dans la base de données KLEMS?

  Il existe plusieurs estimateurs qui sont insensibles (ou moins sensibles) aux observations atypiques, dont les estimateurs par les moindres carrés médians, les estimateurs M, les estimateurs par les moindres carrés tronqués et les estimateurs S. Ces estimateurs utilisent des fonctions des données qui, de diverses façons, tiennent compte de l'incidence des observations atypiques. Dans le présent document, l'estimateur S est utilisé parce qu'il est insensible aux observations atypiques dans les variables dépendantes et indépendantes.

- Comment les observations atypiques influent-elles sur les estimations MCO, et cette incidence est-elle importante d'un point de vue économique?

Par rapport à l'estimateur S, l'estimateur MCO sous-estime la croissance de la PTF de 0,35 point de pourcentage en moyenne, et de 4,3 points de pourcentage pour certaines industries. Des écarts de cette ampleur ne sont pas négligeables, surtout lorsqu'ils sont cumulés sur 43 ans.

- Les observations atypiques sont-elles distribuées également d'une année à l'autre?

L'analyse des périodes où des observations atypiques sont présentes montre qu'elles tendent à se regrouper autour de certains événements. Notamment, le premier choc pétrolier du début des années 1970, la récession de 1980-1981 et la récession de 1990 semblent être d'importantes sources d'observations atypiques.

- À quoi les observations atypiques sont-elles attribuables?

Les chocs macroéconomiques représentent la principale source d'observations atypiques. Il semble que lorsque les industries traversent une période de stress, que ce soit en raison d'une augmentation rapide des coûts des facteurs combinée à une demande inélastique, ou encore d'une réduction de la demande globale, la fonction de réponse des industries soit différente de celle observée au cours des périodes d'expansion « normales ». Les réponses varient grandement d'une industrie à l'autre et, aux fins de l'estimation de la PTF, représentent des ajustements qui se situent probablement au-delà de la portée des hypothèses traditionnelles concernant la croissance de la productivité.

- Quelle est l'importance des observations atypiques dans l'ensemble de données KLEMS?

Les observations atypiques constituent une caractéristique importante de l'ensemble de données KLEMS. Dans le présent document, où une fonction de production de Cobb-Douglas est utilisée, il s'avère que jusqu'à 21 % des observations de l'échantillon sont atypiques d'une certaine façon. Ne pas tenir compte de ces observations peut influer sur les estimations paramétriques et se traduire par une inférence erronée. Les chercheurs qui utilisent l'ensemble de données KLEMS doivent se demander s'ils devraient employer des techniques spéciales afin de tenir compte des observations atypiques.

1. *Introduction*

Grâce à l'expansion de la base de données KLEMS (capital, travail, énergie, matières et services) de Statistique Canada, les chercheurs disposent d'un riche ensemble de données pour examiner les relations économiques. Cette base de données a fait l'objet d'une extrapolation rétrospective de 1997 à 1961 à l'aide des classifications définies dans le Système de classification des industries de l'Amérique du Nord de 1997 (SCIAN 1997). Auparavant, elle renfermait des agrégations selon la Classification type des industries de 1980 (CTI 1980) de 1961 à 1997, et selon le SCIAN 1997 de 1997 à 2003.

La base de données élargie couvre maintenant 43 ans, de 1961 à 2003, et fournit davantage de points de données à partir desquels des inférences peuvent être faites. Elle renferme des valeurs nominales et des valeurs réelles, ainsi que les indices de prix connexes, pour la production brute, le produit intérieur brut, les services du capital, les services du travail et les facteurs intermédiaires. Des estimations de la productivité du travail et de la productivité multifactorielle sont également incluses.

Les données désagrégées permettent aux économistes et aux décideurs de mieux comprendre le comportement et l'évolution des industries au Canada. Notamment, l'ensemble élargi de données désagrégées constitue une riche source de renseignements à partir de laquelle on peut produire un ensemble à jour d'estimations paramétriques de la productivité totale des facteurs (PTF) par industrie et examiner d'autres relations économiques.

Toutefois, les séries chronologiques incluses dans l'ensemble de données couvrent un grand nombre d'industries susceptibles de subir avec le temps divers chocs économiques et autres. En raison de facteurs tels que les changements apportés à la méthodologie, l'erreur de mesure ainsi que les chocs économiques agrégés et idiosyncratiques, il peut être difficile de faire des inférences à partir des données. En outre, comme certaines industries peuvent être plus sensibles que d'autres aux événements économiques, il peut y avoir du bruit dans les données au niveau de l'industrie, surtout lorsqu'on les compare à des séries chronologiques agrégées.

La base de données KLEMS est établie notamment à partir des tableaux des ressources et des emplois des Comptes nationaux, dont la préparation nécessite le rapprochement de nombreuses sources de données. Les classifications des industries et des produits changent avec le temps, et les séries doivent être raccordées. De nouvelles sources d'information (p. ex., sur les prix) deviennent disponibles et doivent être intégrées dans les tableaux. Malgré toutes les mesures prises afin d'assurer la continuité des séries, la présence de points irréguliers peut faire en sorte que la qualité des séries ne soit pas idéale pour certaines fins.

Les chocs économiques idiosyncratiques et agrégés influent aussi sur les données KLEMS. Les chocs idiosyncratiques s'entendent des événements économiques qui touchent une industrie ou un secteur donné, mais non l'ensemble du secteur des entreprises. Ces chocs peuvent être causés, par exemple, par des progrès technologiques, par des changements touchant les obstacles tarifaires ou non tarifaires, les taux d'imposition ou de subvention ou les préférences des consommateurs, ou encore par une concurrence étrangère accrue dans une industrie donnée. Les chocs économiques agrégés, en revanche, sont des événements qui touchent la majorité, voire l'ensemble des industries au Canada.

Le présent document montre progressivement que les observations inhabituelles, qui sont appelées observations atypiques étant donné la multitude de causes possibles, constituent un élément important des processus de production de données sur les industries. C'est pourquoi il est important d'examiner les données sous-jacentes afin de comprendre comment ces caractéristiques des données peuvent influer sur l'estimation et, partant, sur l'inférence[1]. Bien que ce soit le cas pour n'importe quel ensemble de données, c'est particulièrement important lorsque la base de données KLEMS est utilisée. Les données sous-jacentes sont bruitées et soumises à des chocs économiques et autres. Ces problèmes font en sorte qu'il est difficile de déterminer les « vraies » estimations paramétriques sous-jacentes en se servant de techniques d'estimation courantes telles que la méthode des moindres carrés ordinaires (MCO).

Les problèmes occasionnés par les observations atypiques peuvent, dans certains contextes, être surmontés en adoptant une stratégie d'estimation différente, comme il est démontré ici. Pour illustrer ce point, une comparaison est établie entre un estimateur MCO et un estimateur qui est moins sensible aux observations atypiques. L'écart entre les deux ensembles d'estimations montre, *ceteris paribus*, comment les observations atypiques influent sur les estimations MCO. Ce constat ne signifie pas que les estimations MCO sont inexactes; il laisse plutôt entendre que les chercheurs, avant de procéder à l'estimation, devraient prendre le temps d'évaluer si la période d'échantillonnage convient pour l'examen d'une relation économique donnée.

Bien que les observations atypiques puissent avoir une forte incidence sur les estimations paramétriques et l'inférence, les données ne sont pas de mauvaise qualité pour autant. Il faut toutefois que les chercheurs qui utilisent ces données sachent comment les observations atypiques se manifestent. Bien souvent, celles-ci indiquent qu'un événement hors de l'ordinaire est survenu au sein d'une industrie donnée au cours d'une année donnée. La présence de telles observations ne doit pas servir à conclure que les données sont inexactes ou inutilisables.

Le présent document propose un examen de la façon empirique de traiter les observations atypiques lorsqu'on estime la productivité totale des facteurs (PTF) à partir d'une fonction de production simple. La section 2 renferme une description des données KLEMS utilisées dans le processus d'estimation. Le modèle économique employé est décrit à la section 3, alors que la section 4 traite des tests de racine unitaire et illustre les difficultés que pose pour les praticiens l'utilisation de données bruitées désagrégées. Bien qu'il ne s'agisse pas d'un traité sur les tests de racine unitaire, cette section est importante pour les lecteurs désireux de vérifier l'hypothèse de racine unitaire avec les données KLEMS, première étape importante dans tout exercice économétrique utilisant des données chronologiques. À la section 5, il est question du modèle économétrique qui est utilisé pour établir les estimations paramétriques de la PTF. Les méthodes d'estimation sont ensuite décrites à la section 6. Comme les données peuvent comporter des observations atypiques, la méthode des moindres carrés ordinaires (MCO) n'est pas appropriée. Un estimateur S est donc utilisé pour fournir des estimations robustes et les juxtaposer aux estimations MCO. Enfin, un examen des estimations est présenté à la section 7, puis la section 8 conclut l'analyse.

1. Le terme *observation atypique* est utilisé dans le présent document pour désigner une catégorie d'observations qui sont sensiblement différentes des autres observations incluses dans l'échantillon.

2. *Données KLEMS*

La présente étude utilise les données sur le produit intérieur brut (PIB) réel, les services du capital et les services du travail qui sont incluses dans la base de données KLEMS (capital, travail, énergie, matières et services) de Statistique Canada. L'ensemble de données comprend 88 industries du secteur des entreprises et une industrie du secteur non commercial qui correspondent aux regroupements d'industries du niveau P du SCIAN 1997[2].

Pour chaque industrie, le PIB réel est tiré directement des tableaux des entrées-sorties produits par la Division des comptes des industries de Statistique Canada.

Les services du capital constituent une mesure de la contribution du stock de capital à la valeur ajoutée intérieure calculée pour les Comptes canadiens de productivité. Pour chaque industrie, on utilise les données sur les investissements et on applique la méthode de l'inventaire permanent afin d'établir les stocks de capital pour une gamme d'actifs. Les services du capital pour chaque industrie sont ensuite calculés en faisant la moyenne pondérée des stocks individuels employés. Le coût relatif du capital est utilisé pour pondérer les divers actifs. Des taux de dépréciation différents selon les actifs, qui influent sur le coût du capital, garantissent que les actifs qui perdent de la valeur rapidement génèrent des valeurs de service plus élevées pour couvrir les coûts (voir Baldwin et Gu, à venir; Schreyer, Diewert et Harrison, 2005; Gellatly, Tanguay et Yan, 2002; et Harchaoui et Tarkhani, 2003).

Les services du travail représentent une mesure de la contribution du facteur travail à la valeur ajoutée intérieure calculée pour les Comptes canadiens de productivité. Généralement, dans les études économétriques, ce sont les heures travaillées qui sont utilisées comme mesure du facteur travail. Toutefois, cette méthode ne tient pas compte des différences, entre les industries et les professions, au chapitre des compétences et du capital humain. Les estimations des services du travail calculées pour les comptes de la productivité utilisent des taux de rémunération relatifs pour agréger les heures travaillées afin d'établir une série du facteur travail pour chaque industrie. Les séries des services du travail supposent que les travailleurs sont rémunérés pour leur produit marginal. Par conséquent, les travailleurs qui touchent une rémunération élevée contribuent plus que les autres à la production pour chaque heure travaillée (Gu et coll., 2002).

2. Il existe aussi des données pour trois autres regroupements : secteur des entreprises, niveau M et niveau S.

3. *Spécification du modèle*

La productivité totale des facteurs (PTF) mesure la contribution à la croissance économique qui n'est pas attribuable à des facteurs tangibles. Elle est souvent considérée comme une mesure des progrès techniques non incorporés sous-jacents. Cependant, elle peut être perçue également comme saisissant une vaste gamme de facteurs, dont l'erreur de mesure ou des variables qui peuvent être incluses dans une fonction de production mais qui sont difficiles à mesurer, telles que les changements dans la structure de gestion, les externalités de réseau ou la contribution de la recherche-développement ou de l'infrastructure publique à la production.

La PTF peut être estimée paramétriquement ou non paramétriquement, les deux approches reposant sur l'hypothèse qu'il existe une fonction de production industrielle. La méthode non paramétrique consiste généralement à estimer les parts du travail et du capital dans le revenu, puis à utiliser ces parts comme facteurs de pondération dans une fonction de production où une hypothèse de production à rendements constants à l'échelle est imposée. À l'aide des estimations des facteurs capital et travail, mises à l'échelle par leurs poids correspondants, on calcule une estimation de la croissance que devrait connaître le produit intérieur brut (PIB) étant donné la variation des facteurs. Invariablement, cette estimation est différente de la croissance réelle du PIB, et l'écart est désigné comme étant la croissance de la PTF (voir, par exemple, Jorgenson et Stiroh, 2000; Jorgenson, Gollop et Fraumeni, 1987).

Les estimations paramétriques sont fondées sur une hypothèse de forme fonctionnelle pour la fonction de production du PIB. Les facteurs capital et travail sont utilisés pour produire économétriquement des estimations de leurs élasticités. Les estimations de la PTF sont toutefois dépendantes de la façon dont les données sont employées. Si on utilise les niveaux logarithmiques des données, alors une tendance temporelle peut être employée pour saisir la PTF. Toutefois, s'il s'agit de données en différences logarithmiques, la tendance temporelle dans le modèle en niveaux devient constante, de sorte que l'ordonnée à l'origine dans le modèle en différences saisit la croissance de la PTF. La décision quant à l'utilisation des niveaux logarithmiques ou des différences logarithmiques dépend finalement de la nature temporelle des données.

Dans la présente étude, une fonction de production Cobb-Douglas est utilisée pour établir une estimation paramétrique de la PTF :

$$\ln PIB_t = PTF(t) + \beta_K \ln K_t + \beta_L \ln L_t \qquad (1)$$

où PIB est le PIB réel au temps t, K représente les services du capital et L correspond aux services du travail. Les paramètres β_K et β_L saisissent l'élasticité du PIB en ce qui touche les service du capital et les services du travail respectivement. Ils sont interprétés comme les élasticités des services du capital et des services du travail. Deux formes du modèle sont utilisées. La première n'applique aucune contrainte aux paramètres, alors que la seconde impose des rendements constants à l'échelle sur la fonction de production :

$$\beta_K + \beta_L = 1. \qquad (2)$$

La façon dont la PTF est estimée (comme une tendance temporelle utilisant les niveaux logarithmiques plutôt qu'une constante utilisant les différences logarithmiques) dépend de la forme fonctionnelle de $PTF(t)$. À priori, la forme fonctionnelle que les données permettront est incertaine, et le choix sera fait selon que l'on considère que les données renferment une racine unitaire ou une tendance déterministe.

4. *Tests de racine unitaire*

La présente section n'est pas un examen exhaustif des questions touchant l'analyse des séries chronologiques ou les tests de racine unitaire. Elle donne plutôt un aperçu des raisons pour lesquelles il est important de tenir compte des racines unitaires et de la façon dont les observations atypiques et les changements structurels peuvent influer sur les tests de racine unitaire appliqués aux séries tirées de la base de données KLEMS (capital, travail, énergie, matières et services)[3].

Lorsque des séries économiques affichent une augmentation ou une diminution stochastique dans le temps, il est possible de produire des estimations de coefficient statistiquement significatives à l'aide d'une analyse multivariée uniquement parce que les séries utilisées pour ce faire présentent une tendance. Cette caractéristique de l'analyse chronologique, qui a été découverte par Yule (1926) et reprise plus tard par Granger et Newbold (1974), a pour nom « régression illusoire ». Avec des résultats tirés d'une régression illusoire, il faut se fier à la série tendancielle, et non à une relation économique sous-jacente, pour établir des estimations paramétriques qui soient statistiquement significatives. La régression illusoire est généralement associée à des séries qui suivent un processus de racine unitaire et force les praticiens des séries chronologiques à faire attention lorsqu'ils examinent ces types de relations. C'est pourquoi une grande attention est portée aux tests de racine unitaire.

On dit qu'une série suit une racine unitaire lorsque les innovations dans la série sont permanentes. Les innovations sont permanentes lorsque le paramètre autorégressif dans la relation univariée entre Y_{t-1} et Y_t est égal à 1. Cette série diffère d'une série qui suit une tendance déterministe dans le temps, où les innovations créent des divergences temporaires par rapport à la tendance.

Lorsque les praticiens des séries chronologiques analysent une relation multivariée, ils doivent premièrement déterminer le type de tendance présente dans les données, problème que viennent compliquer deux facteurs :

1. Sous l'hypothèse nulle qu'une série chronologique suit un processus de racine unitaire, la statistique de test peut converger vers une distribution limite non uniforme.
2. La puissance des tests de racine unitaire est souvent peu élevée.

3. Pour en apprendre davantage sur la question, voir Maddala et Kim (1998).

La première complication est surmontée grâce à des techniques de simulation qui fournissent les valeurs critiques. Toutefois, la seconde complication demeure. Les tests de racine unitaire tentent de trouver les changements permanents dans les données chronologiques. Lorsque des événements ponctuels, telle une récession, influent sur le niveau de la série, ils peuvent mener à une inférence erronée à partir des tests de racine unitaire (Perron, 1989; Madalla et Kim, 1998). Autrement dit, des tests couramment appliqués peuvent laisser supposer à tort la présence d'une racine unitaire. Il revient alors à l'analyste de se demander si ces événements ponctuels entraînent des ruptures structurelles dont il faut tenir compte. Il s'agit là d'un problème sérieux lorsque les données KLEMS sont analysées.

Deuxièmement, les praticiens doivent déterminer comment traiter les données, en se fondant sur le résultat des tests de racine unitaire. Si les séries en question suivent une racine unitaire, il faut souvent faire en plus un test de cointégration, qui consiste à déterminer si une combinaison linéaire de processus qui suivent un processus de racine unitaire est stationnaire ou non[4]. Si c'est le cas, il est quand même possible d'utiliser les niveaux logarithmiques des données pour établir des estimations de la relation à long terme entre les variables. Si les séries en question suivent une tendance déterministe, alors l'inclusion de variables déterministes se traduit par des résultats de régression valides. Cependant, si les séries ne sont pas cointégrées, ou si certaines d'entre elles semblent stationnaires alors que d'autres semblent suivre un processus de racine unitaire, l'analyse de régression utilisant les niveaux logarithmiques des données peut se solder par des résultats erronés.

4.1 Observations atypiques et tests de racine unitaire

Les praticiens des séries chronologiques, qu'ils s'intéressent aux prévisions ou aux estimations paramétriques, incluent invariablement un examen de la stationnarité dans leur analyse. En général, si une série présente une tendance, un test de racine unitaire est effectué et une statistique de test est comparée à un ensemble de valeurs critiques. Bien souvent, un second test de racine unitaire est effectué, et le résultat comparé à celui du test initial. Il s'agit d'une analyse confirmative (Maddala et Kim, 1998). Selon les résultats du ou des tests de racine unitaire, soit les données sont utilisées telles quelles, soit elles sont transformées pour tenir compte de la non-stationnarité ou soit, dans le cas d'une analyse multivariée, un prétest de cointégration est effectué.

Au cours de ce processus, il arrive que les praticiens prêtent peu d'attention aux données sous-jacentes avant la tenue des tests de racine unitaire, ce qui peut entraîner de graves problèmes si des observations atypiques ou des changements structurels sont présents dans les séries chronologiques. La puissance des tests de racine unitaire est souvent peu élevée, et ceux-ci ne parviennent pas, bien souvent, à faire la distinction entre, d'une part, les processus de racine unitaire et, d'autre part, les ruptures structurelles, les changements de tendance ou d'autres changements dans les processus de production de données sous-jacentes (Perron, 1989; Maddala et Kim, 1998). Des observations atypiques peuvent saisir, ou entourer, ces événements et provenir de diverses sources, que ce soit l'erreur de mesure, les points tournants du cycle économique ou les chocs économiques (Franses, 1998; Chang, Tia et Chen, 1988; et Fox, 1972).

4. De façon plus générale, un test de cointégration détermine si une combinaison linéaire de I(d) variables forme une série qui est I(d-1).

Par conséquent, la présence de ruptures structurelles ou d'observations atypiques peut influer sur les estimations paramétriques et se traduire par une piètre inférence quant à la présence de processus de racine unitaire.

Par conséquent, il est important que les chercheurs examinent leurs données afin de repérer les observations atypiques ou les changements dans les processus sous-jacents de production de données qui, pour une raison ou une autre, peuvent avoir une influence indue sur les paramètres de régression utilisés pour les prétests. En outre, lorsque ces difficultés se présentent, il peut être important d'appliquer un jugement professionnel. Cela ne signifie pas nécessairement qu'une transformation est de mise, mais suggère plutôt que lorsqu'une série qui semble suivre une tendance linéaire (tendance stochastique) génère un résultat de test paramétrique qui laisse supposer que la série suit un processus de racine unitaire (est stationnaire autour d'une tendance), les chercheurs doivent se demander pourquoi une telle contradiction apparaît.

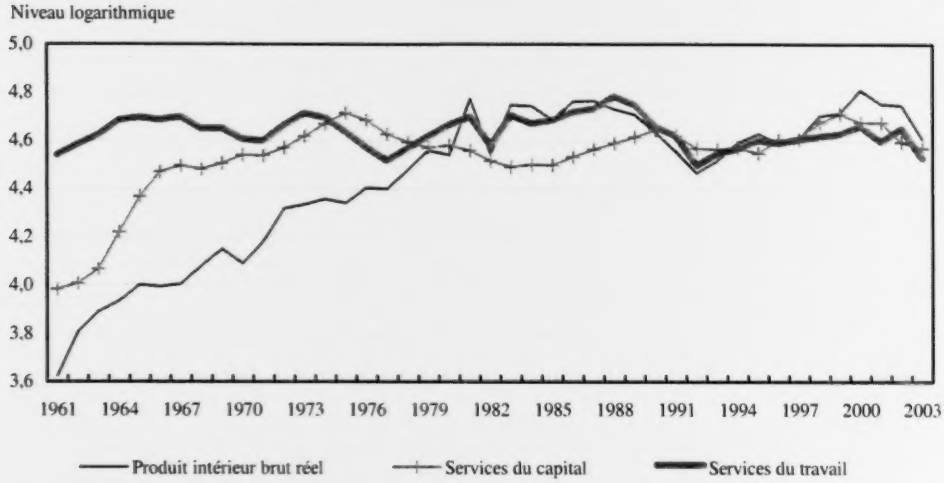
4.2 Tests de racine unitaire des industries échantillonnées

Il est de plus en plus important d'analyser les séries chronologiques lorsque des données désagrégées sont utilisées. Comme les données désagrégées peuvent être plus sensibles aux chocs et afficher des tendances ainsi que des changements de tendance ou de niveau qui sont différents de ceux des séries agrégées, le problème peut s'en trouver amplifié. Dans la base de données KLEMS, de nombreuses industries sont soumises à ces types de chocs, ce qui rend difficile la réalisation de tests de racine unitaire.

Par exemple, les niveaux du PIB réel, des services du capital et des services du travail pour les industries de la fabrication de textiles et de produits textiles, de la fabrication de matériel ferroviaire roulant et de la fabrication de matériel informatique et périphérique sont présentés dans les figures 1 à 3. Dans l'industrie de la fabrication de textiles et de produits textiles, on observe une hausse du PIB réel jusqu'à la récession de 1981, après quoi le niveau du PIB se stabilise. Le niveau des services du capital augmente rapidement au début de la période, avant de plafonner vers la fin des années 1960 pour demeurer assez constant par la suite. Le niveau des services du travail semble quant à lui fluctuer autour d'une tendance.

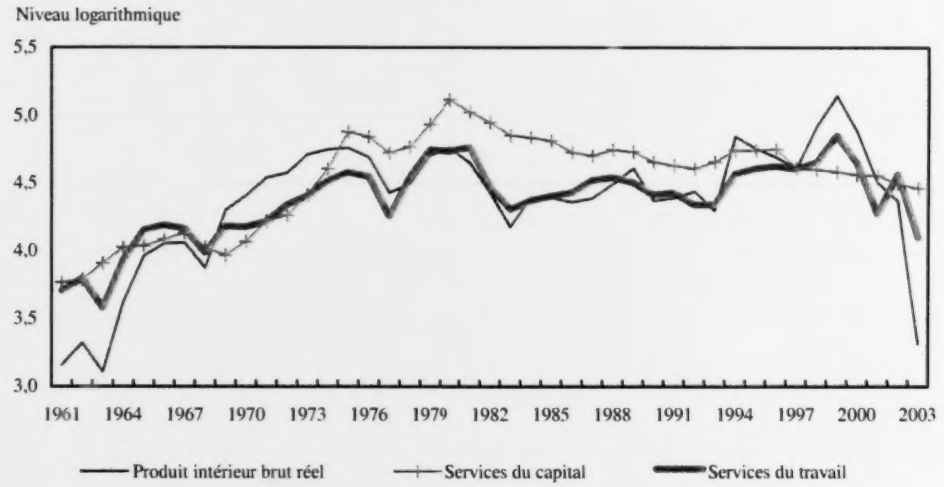
Dans l'industrie de la fabrication de matériel ferroviaire roulant, le PIB réel et les services du travail subissent à un certain nombre de points des variations de niveau qui semblent permanentes. Ainsi, on observe une hausse des niveaux du PIB réel et des services du travail au début de la période, vers la récession de 1981 et à la fin de la période. Le niveau des services du capital augmente de 1961 à 1980, puis il commence à diminuer de 1981 jusqu'à la fin de la période.

Figure 1
Fabrication de textiles et de produits textiles



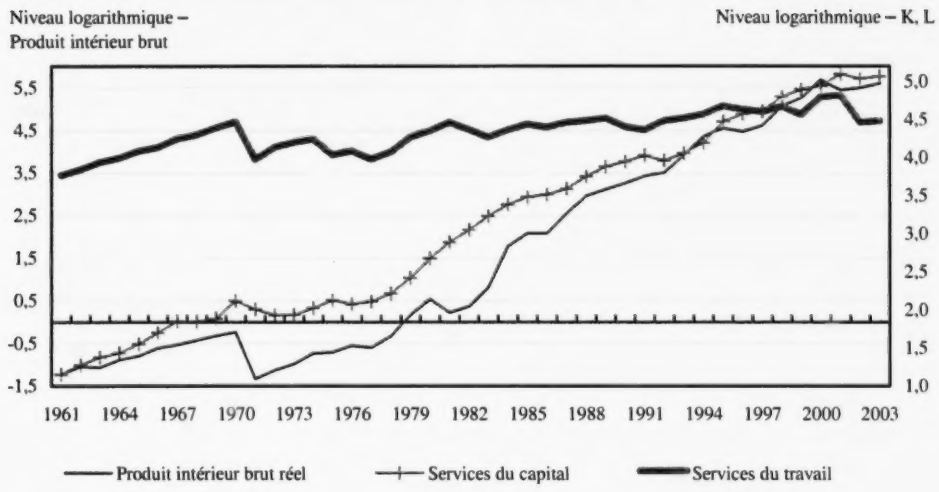
Source : Statistique Canada.

Figure 2
Fabrication de matériel ferroviaire roulant



Source : Statistique Canada.

Figure 3
Fabrication de matériel informatique et périphérique



Source : Statistique Canada.

Dans l'industrie de la fabrication de matériel informatique et périphérique, le niveau des services du travail augmente dans le temps. Toutefois, on observe une baisse qui semble permanente en 1971, après quoi le niveau des services du travail devient plus instable. Le niveau du PIB réel suit les mouvements de celui des services du travail au début de la période. Après la baisse de 1971, il augmente rapidement à mesure que la hausse du niveau des services du capital commence à s'accélérer. Les services du capital augmentent plus rapidement que les services du travail, et ils subissent une moins forte baisse au début des années 1970. Après le recul du début des années 1970, les services du capital augmentent pour la plus grande partie de la période d'échantillonnage.

Les figures 1 à 3 illustrent la difficulté que pose l'ensemble de données désagrégées KLEMS. Les données sont fluctuantes, certaines séries semblent suivre des processus de racine unitaire au cours de la période d'échantillonnage, alors que d'autres semblent suivre des tendances linéaires. D'autres encore semblent contenir des ruptures structurelles dans leurs niveaux ou leurs tendances. En présence de ces types de chocs, les tests de racine unitaire traditionnels peuvent donner de piètres résultats.

Pour fins de démonstration, deux ensembles de tests de racine unitaire / stationnarité sont appliqués aux séries des figures 1 à 3. Le premier est le test de Dickey-Fuller amélioré (DFA) (tableau 1). Pour la majorité des séries, le test DFA ne parvient pas à rejeter l'hypothèse nulle qu'une racine unitaire est présente. L'hypothèse nulle n'est rejetée que pour le PIB, les services du capital et les services du travail de la série de la fabrication de textiles et de produits textiles, et pour les services du travail de la série de la fabrication de matériel informatique et périphérique.

Tableau 1
Tests DFA[1] de racine unitaire

	Ho : La série contient une racine unitaire		
	Fabrication de textiles et de produits textiles	Fabrication de matériel ferroviaire roulant	Fabrication de matériel informatique et périphérique
ln PIB[2]	-2,99*	-2,27	0,53
ln L	-3,00*	-2,80*	-2,55
ln K	-4,69*	-2,33	-0,04

1. Dickey-Fuller amélioré.
2. Produit intérieur brut.
Nota : L'astérisque (*) indique le rejet de l'hypothèse nulle au seuil de signification de 10 %.
Source : Statistique Canada.

Bien que les résultats semblent plausibles pour la série de la fabrication de textiles et de produits textiles, ils ne sont pas complètement conformes aux attentes antérieures pour toutes les autres séries. Notamment, dans l'industrie de la fabrication de matériel informatique et périphérique, la série du PIB réel peut être décrite comme étant constituée de trois segments, allant de 1961 au début des années 1970, du milieu des années 1970 au début des années 1980, et enfin du milieu des années 1980 à la fin de la période d'échantillonnage. Au cours des deux premières sous-périodes, les données semblent suivre une tendance linéaire qui est interrompue par une transition rapide causée par des événements économiques dont on sait qu'ils ont eu de vastes répercussions. Il en va de même pour la série des services du capital pour l'industrie de la fabrication de matériel ferroviaire roulant, qui augmente jusqu'à la récession des années 1980-1981 et décline par la suite. Dans ces cas, un processus de racine unitaire peut être erronément appliqué à des séries qui subissent un changement structurel attribuable à des événements économiques précis.

Comme il est difficile de tester statistiquement l'hypothèse de racine unitaire, les praticiens cherchent souvent une confirmation de leurs résultats en effectuant un test de racine unitaire supplémentaire (voir Madalla et Kim, 2003, p. 126 à 128). Si les résultats de ce second test et du test initial sont identiques, ils sont jugés plus crédibles. La confirmation est vue comme un élément additionnel prouvant que la série suit ou non un processus de racine unitaire. Toutefois, le second test peut contredire le premier test, ce qui est souvent le cas. Les options disponibles sont alors limitées.

Suivant la veine de l'analyse confirmative, un test KPSS (Kwaitkowski, Phillips, Schmidt et Shin, 1992) est appliqué à chacune des séries du PIB réel, des services du capital et des services du travail. Sous l'hypothèse nulle dans le test KPSS, la série suit une tendance linéaire.

Les résultats sont présentés au tableau 2. Les tests KPSS suggèrent que toutes les séries, à l'exception des séries des services du travail pour la fabrication de textiles et de produits textiles et la fabrication de matériel informatique et périphérique, suivent un processus de racine unitaire. Ces résultats diffèrent de ceux des tests DFA pour le PIB et les services du capital de la fabrication de textiles et de produits textiles, pour les services du travail de la fabrication de matériel ferroviaire roulant et pour les services du travail de la fabrication de matériel informatique et périphérique. Pour toutes les autres séries chronologiques, les résultats des tests semblent se confirmer les uns les autres, renforçant la notion que ces séries renferment une racine unitaire.

Tableau 2
Test KPSS[1] de stationnarité

	Ho : La série est stationnaire		
	Fabrication de textiles et de produits textiles	Fabrication de matériel ferroviaire roulant	Fabrication de matériel informatique et périphérique
ln PIB[2]	0,20*	0,16*	0,17*
ln L	0,08	0,19*	0,06
ln K	0,15*	0,20*	0,12*

1. Kwaitkowski, Phillips, Schmidt et Shin.
2. Produit intérieur brut.
Nota : L'astérisque (*) indique le rejet de l'hypothèse nulle au seuil de signification de 10 %.
Source : Statistique Canada.

Les résultats obtenus sous-entendent qu'un mélange de tendances déterministes et de tendances stochastiques sont présentes. Si ces résultats sont considérés comme étant crédibles, alors tout modèle qui est cohérent d'une industrie à l'autre devra rapprocher la nature apparemment divergente des séries de facteurs et de réponses. Cependant, un examen plus approfondi des données révélera que la faible puissance des tests de racine unitaire conduit probablement à des résultats erronés, ce qui rend l'inférence difficile. Comme Perron (1989) l'a démontré, lorsqu'il y a une rupture structurelle ponctuelle, ou un changement ponctuel dans la tendance d'une série chronologique qui est stationnaire autour de cette tendance, les tests de racine unitaire peuvent indiquer à tort la présence d'une racine unitaire.

4.3 Examen détaillé de l'industrie de la fabrication de textiles et de produits textiles

L'examen visuel des séries chronologiques des figures 1 à 3 fournit un exemple qui est similaire dans sa structure à l'argument présenté par Perron (1989). Le PIB de la fabrication de textiles et de produits textiles semble suivre une tendance linéaire qui se modifie vers 1980, ce qui correspond à un des exemples examinés par Perron. Par conséquent, l'hypothèse de stationnarité pour cette série est examinée de plus près. Notamment, le résultat du test KPSS est examiné de façon plus détaillée parce qu'il fournit l'occasion de démontrer succinctement l'incidence du changement de tendance.

La statistique de test KPSS est un test de la constance des paramètres. Elle découle du modèle économétrique où :

$$y_t = \alpha_t + \beta t + e_t$$
$$\alpha_t = \alpha_{t-1} + u_t$$

et $u_t \sim iid(0, \sigma_u^2)$. Lorsque $\sigma_u^2 = 0$, la série suit une tendance linéaire, alors que $\sigma_u^2 > 0$ suppose que le paramètre suit une marche aléatoire[5]. La statistique de test est formée au moyen des

5. Cette description du test est une version résumée de la description donnée par Maddala et Kim (1998). Pour de plus amples détails, voir Maddala et Kim (1998).

résidus de la régression $\ln PIB_t = \alpha + \beta t + \varepsilon_t$ où t est une tendance linéaire (modèle 1). Les résidus du modèle 1 sont utilisés pour créer une série de sommes partielles :

$$S_t = \sum_{t=1}^{n} e_t \quad n = 1 \ldots T.$$

Si la série suit une tendance linéaire, il ne devrait pas y avoir de divergence permanente entre la série chronologique réelle et le modèle ajusté. Par conséquent, toutes les sommes partielles des résidus devraient être près de 0. Pour former la statistique de test, il convient de prendre note que les sommes partielles peuvent être exprimées sous la forme :

$$S_t = S_{t-1} + e_t .$$

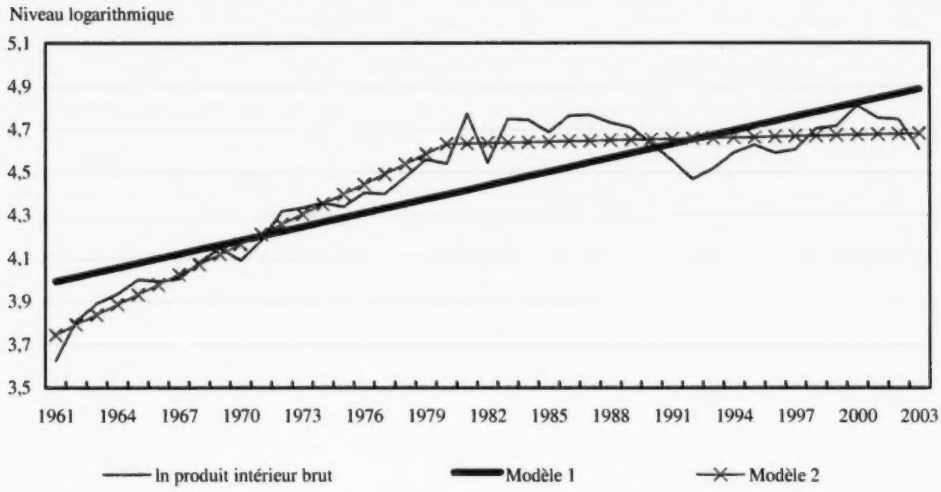
Lorsqu'une racine unitaire est présente, des changements permanents dans le niveau de la série se traduiront par des sommes partielles dont la variance n'est pas nulle. La statistique de test est donc obtenue en combinant une estimation de la variance des sommes partielles avec une estimation de la variance à long terme des résidus estimés. Les valeurs critiques asymptotiques sont fournies dans Kwaitkowski et coll. (1992). Le test est fourni dans un certain nombre de logiciels statistiques, qui fournissent également les valeurs critiques asymptotiques. (Pour plus de renseignements, voir : Kwaitkowski et coll., 1992; Franses, 1998; Madalla et Kim, 1998).

Cependant, une erreur de spécification du modèle nul peut également se traduire par une série de sommes partielles résiduelles dont la variance n'est pas nulle. Bien que le modèle 1 (modèle nul KPSS) au tableau 2 suggère la présence d'une racine unitaire dans le PIB de la fabrication de textiles et de produits textiles, la figure 4 suggère qu'il y a un changement dans la tendance de la série vers 1980. Si le modèle nul (modèle 1) comporte une erreur de spécification, cela aura une incidence sur les résidus et le test. Alors, plutôt que d'utiliser un modèle nul qui présume la même tendance pour l'ensemble de l'échantillon, supposons que le modèle nul permet un changement dans la tendance en 1980. Plus particulièrement, prenons le modèle nul (modèle 2) $\ln PIB_t = \alpha + \beta_1 t + \delta_2 \beta_2 (t - t_L) + \varepsilon_t$ où t et t_L sont des tendances temporelles pour l'ensemble de la période et pour la période postérieure à 1980 respectivement, et δ_2 est une variable indicatrice dont la valeur est 0 de 1961 à 1979 et 1 de 1980 à 2003[6].

Les valeurs ajustées des deux modèles sont représentées graphiquement en fonction du log du PIB des textiles à la figure 4, qui montre clairement que les valeurs ajustées du modèle 2 suivent le PIB réel mieux que celles du modèle 1. En outre, les sommes partielles des résidus du modèle 1 et du modèle 2 sont manifestement différentes (figure 5). Selon le modèle 1, la variance de la série des sommes partielles semble supérieure à 0, ce qui laisse entendre que la série n'est pas stationnaire et mène au résultat du test dans le tableau 2. Toutefois, la variance de la série des sommes partielles dans le modèle 2 semble près de 0, ce qui signifie que la série est stationnaire autour d'une tendance comportant une rupture. La statistique de test KPSS selon le modèle 2 se chiffre à 0,05 et n'est donc pas significative au seuil de 10 %, ce qui signifie que la série est stationnaire autour d'une tendance.

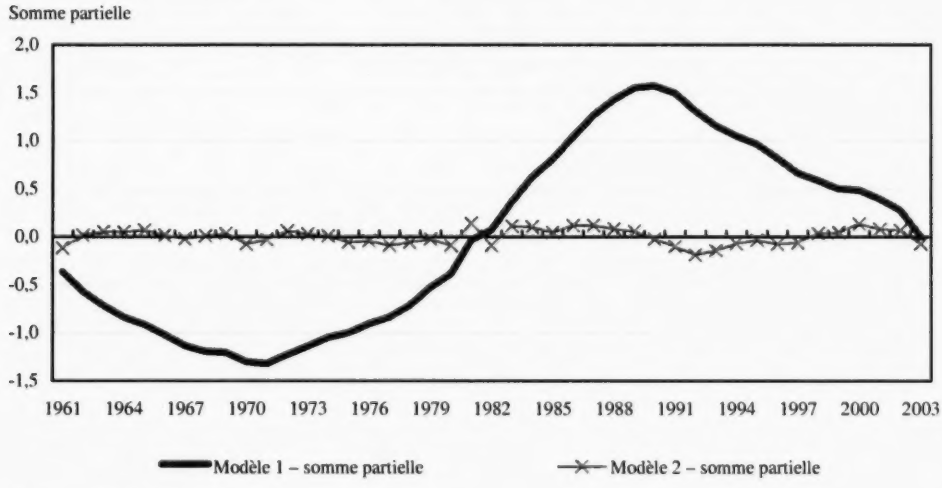
6. Il s'agit du modèle B de Perron (1989).

Figure 4
Modèles nuls pour le PIB[1] de la fabrication de textiles et de produits textiles



1. Produit intérieur brut.
Source : Statistique Canada.

Figure 5
Sommes partielles pour les erreurs des modèles nuls



Source : Statistique Canada.

Les résultats divergents du modèle 1 et du modèle 2 suggèrent que c'est l'erreur de spécification du modèle nul 1, plutôt que la présence d'une racine unitaire, qui est le facteur dominant dans le rejet de l'hypothèse de présence d'une tendance au tableau 2. De tels résultats devraient donner aux chercheurs des raisons de prendre le temps d'étudier leurs résultats lorsqu'ils utilisent la base de données KLEMS.

L'examen du PIB des textiles suffit à démontrer que les chercheurs qui utilisent la base de données KLEMS doivent faire preuve de prudence. La documentation portant sur les racines unitaires montre que les tests de racine unitaire qui sont appliqués couramment sont peu puissants et qu'ils sont influencés par les changements de tendance et les variations de niveau (Perron, 1989; Maddala et Kim, 1998). Comme des types de variations sont, entre autres, présents dans la base de données KLEMS, il sera difficile d'appliquer simplement ces tests de racine unitaire.

4.4 Tests de racine unitaire sur données de panel

Jusqu'à maintenant, un échantillon d'industries ont été analysées isolément. Toutefois, la base de données KLEMS renferme un panel d'industries. Si les chercheurs tentent de déterminer si chaque série contient une racine unitaire, ou quelles sont les industries qui sont cointégrées, ils seront entraînés dans un processus long et compliqué qui, à la fin, ne leur permettra pas de tirer des conclusions générales au sujet de l'économie. Au contraire, ils seront obligés d'établir un modèle pour chaque industrie séparément et de formuler des hypothèses quant aux points de rupture afin d'accroître la puissance de leurs tests. Cette approche présente de multiples dangers et expose les chercheurs à se faire reprocher de se livrer au forage de données.

Cependant, si toutes les industries incluses dans le panel sont traitées comme ayant la même structure, il est possible d'appliquer les tests de racine unitaire à l'ensemble du panel simultanément. À cette fin, on dispose de tests de racine unitaire basés sur le test DFA et le test KPSS. Ces derniers tentent d'accroître la puissance des tests en exploitant l'ensemble des industries. De fait, les statistiques de test de chaque industrie sont combinées pour former une statistique de test de panel qui est plus puissante que les statistiques de test prises individuellement.

Deux tests de racine unitaire sur données de panel sont menés ici, soit le test IPS (Im, Pesaran et Shin, 1995) et le test LM de Hadri (Hadri, 2000). Le test IPS regroupe les tests DFA effectués pour chaque industrie. Sous l'hypothèse nulle, toutes les séries contiennent une racine unitaire, alors que sous l'hypothèse alternative, au moins une série est stationnaire. Le test LM de Hadri suppose que toutes les industries sont stationnaires sous l'hypothèse nulle, et non stationnaires sous l'hypothèse alternative. Les résidus d'une régression auxiliaire sur une tendance sont utilisés pour former une statistique de test, qui suit le test KPSS. Les résultats du test IPS sont présentés au tableau 3, et ceux du test LM de Hadri au tableau 4.

Le test IPS étaye l'hypothèse nulle que toutes les séries contiennent une racine unitaire, alors que le test LM de Hadri rejette l'hypothèse nulle que les séries sont stationnaires autour d'une tendance. Les tests se confirment l'un l'autre et laissent entendre que les séries de panel devraient être traitées comme si elles suivent un processus de racine unitaire. Toutefois, comme on l'a mentionné précédemment, les données sont bruitées et peuvent comporter de nombreux types de ruptures structurelles et des observations atypiques. Par conséquent, il est probable que les tests de racine unitaire sur données de panel soient affectés de façon analogue aux tests KPSS et aux tests DFA dont il est question dans la section précédente.

Tableau 3
Tests IPS[1] de racine unitaire sur données de panel

H_0 : Toutes les séries contiennent une racine unitaire
H_A : Au moins une série ne contient pas une racine unitaire

	ln PIB[2]	ln L	ln K
Décalages			
1	-1,26	-1,02	-1,27
2	-1,19	-1,07	-1,11

1. Im, Pesaran et Shin.
2. Produit intérieur brut.
Source : Statistique Canada.

Tableau 4
Test LM de Hadri de racine unitaire sur données de panel

H_0 : Toutes les séries chronologiques sont stationnaires

	ln PIB[1]	ln L	ln K
Erreurs homoscédastiques	114,4*	105,5*	141*
Erreurs hétéroscédastiques	100,9*	94,4*	117,4*
Erreurs sériellement dépendantes	29,4*	26,1*	35,0*

1. Produit intérieur brut.
Nota : L'astérisque (*) indique le rejet de l'hypothèse nulle au seuil de signification de 10 %.
Source : Statistique Canada.

Malheureusement, cette information n'indique pas aux chercheurs quelle est la meilleure façon de traiter les données. L'analyse détaillée des trois industries fait nettement ressortir la présence d'un processus complexe de production des données. Il est également clair que dans certains cas, comme celui du PIB de la fabrication de textiles et de produits textiles, une tendance linéaire comportant une rupture se rapproche assez près du processus de production des données. Dans d'autres cas, comme la fabrication de matériel ferroviaire roulant, on voit moins bien si une approximation linéaire correspond aux données. Par conséquent, la méthode « appropriée » pour estimer la PTF dépendra de la mesure dans laquelle le chercheur est à l'aise avec les résultats des tests et de la mesure dans laquelle toutes les séries suivent le même type de tendance (stochastique ou déterministe).

4.5 Transformation des données

Dans le présent document, les résultats des tests de racine unitaire sont considérés comme suspects. Étant donné l'examen des industries échantillonnées et la nature divergente des processus de production des données d'une industrie à l'autre, il semble risqué d'accepter simplement les résultats des tests de racine unitaire sur données de panel. C'est pourquoi les deux hypothèses suivantes sont formulées :

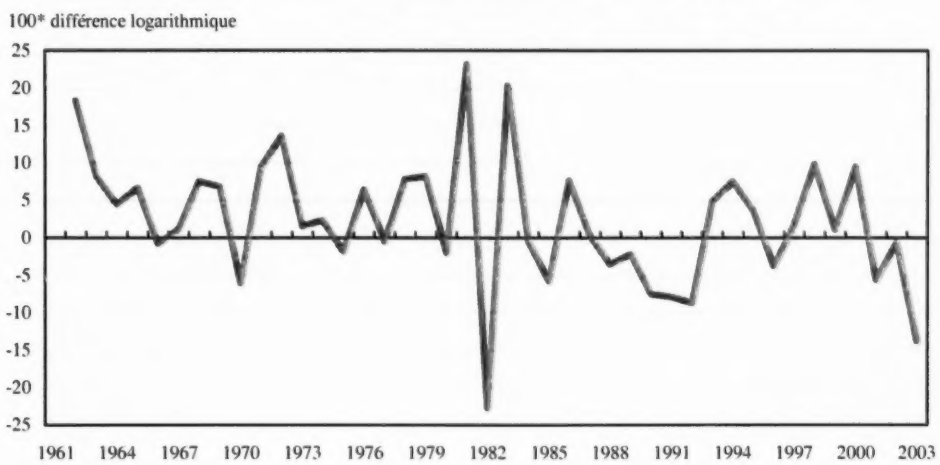
1. Toutes les séries peuvent être rendues covariance-stationnaires en appliquant le filtre de la différence première (1-L) aux données.
2. Les changements contemporains saisiront adéquatement la relation économique entre les variables étudiées.

Il s'agit d'hypothèses audacieuses qui supposent l'absence de problèmes associés aux ruptures structurelles et à la corrélation sérielle des erreurs. Ce sont des problèmes importants qui peuvent influer sur les estimations paramétriques. Toutefois, la présente analyse suppose que ces problèmes sont secondaires en regard de l'incidence des observations atypiques dans l'ensemble de données.

Par exemple, les différences logarithmiques du PIB pour les industries échantillonnées présentent des observations de grandes amplitudes faisant en sorte qu'il est peu probable qu'on puisse identifier avec exactitude les variations moyennes sans d'abord se pencher sur leur influence (figures 4 à 6). Bien que la présente analyse reconnaisse que les variations moyennes et les problèmes de calcul des différences soient importants, elle est axée sur la façon de repérer et de traiter les observations atypiques. Une fois que c'est fait, il devrait être possible de revoir les problèmes touchant les ruptures structurelles et le calcul des différences. Toutefois, pour les fins du présent document, ces problèmes sont exclus du champ d'analyse.

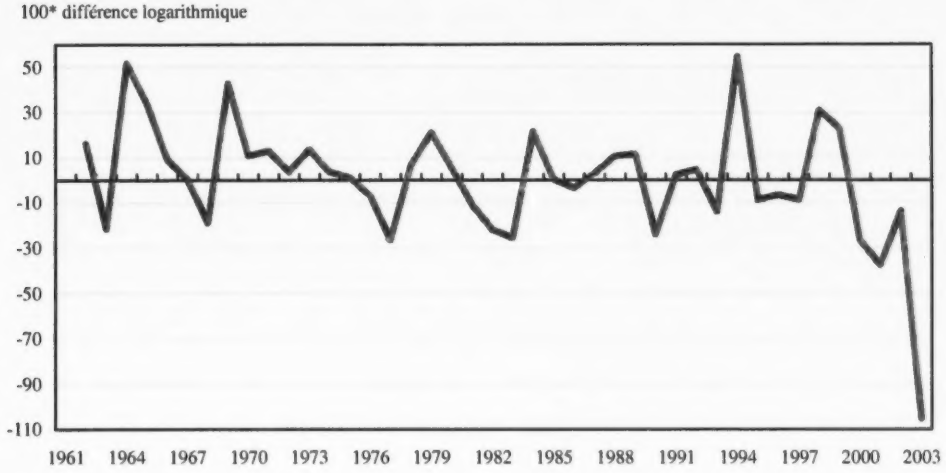
Dans les prochaines sections, l'analyse portera donc sur les données en niveaux logarithmiques qui sont en différence première et multipliées par 100. Pour les petits changements, cela correspondra approximativement au taux de croissance de la série. Cependant, comme le montrent les figures 4 à 6, les données au niveau de l'industrie peuvent être soumises à des changements atypiques d'assez forte amplitude, auquel cas l'approximation du taux de croissance sera inexacte.

Figure 6
Produit intérieur brut de la fabrication de textiles et de produits textiles



Source : Statistique Canada.

Figure 7
Produit intérieur brut de la fabrication de matériel ferroviaire roulant



Source : Statistique Canada.

Figure 8
Produit intérieur brut de la fabrication de matériel informatique et périphérique

Source : Statistique Canada.

Le filtre (1-L) fait en sorte que les données semblent covariance-stationnaires; toutefois, il met également au jour des points de données atypiques qui peuvent avoir une incidence sur certains types d'estimateurs. Quand les données sont en différences logarithmiques, les points tournants du cycle économique, la flambée des prix du pétrole, les chocs idiosyncratiques, les chocs économiques agrégés, l'erreur de mesure et les changements méthodologiques peuvent tous avoir une incidence sur le PIB, les services du capital et les services du travail d'une année à l'autre. Il arrive que ces changements soient d'une grande amplitude par rapport à l'histoire de l'industrie et qu'ils puissent biaiser des estimateurs couramment employés.

5. *Spécification économétrique*

L'analyse préliminaire des données suggère qu'il sera plus fructueux d'utiliser des données en différence première plutôt que de supposer que les séries sont cointégrées ou stationnaires autour d'une tendance linéaire. Une fois que la transformation (1-L) a été appliquée, deux modèles économétriques de base sont utilisés pour établir des estimations de la productivité totale des facteurs (PTF) à l'aide des spécifications Cobb-Douglas de la section 2. Premièrement, (1) devient :

$$\Delta \ln PIB_{i,t} = \alpha_i + \beta_{i,K} \Delta \ln K_{i,t} + \beta_{i,L} \Delta \ln L_{i,t} + \varepsilon_{i,t} \qquad (3)$$

Deuxièmement, après avoir imposé la contrainte des rendements constants à l'échelle (2), l'équation (1) est réduite à :

$$\Delta \ln PT_{i,t} = \alpha_i + \beta_{i,K} \Delta \ln \left(\frac{K_{i,t}}{L_{i,t}} \right) + \varepsilon_{i,t} \qquad (4)$$

où PT est défini comme étant la productivité du travail de telle sorte que :

$$PT_{i,t} = \Delta \ln \left(\frac{PIB_{i,t}}{L_{i,t}} \right).$$

Dans chaque spécification économétrique, la PTF est saisie par une constante, α_i, qui varie d'une industrie à l'autre. Le PIB réel, les services du capital et les services du travail sont indexés dans le temps et selon les industries, où t = 1 … 42 différences logarithmiques de 1962 à 2003 et i = 1 … 88 industries du secteur des entreprises au niveau P d'agrégation dans la base de données KLEMS (capital, travail, énergie, matières et services).

L'interprétation des coefficients des élasticités est analogue à l'interprétation dans le modèle des niveaux logarithmiques décrit à la section 2. Lorsqu'on applique le filtre (1-L), $\beta_{i,K}$ et $\beta_{i,L}$ saisissent l'élasticité des services du capital et des services du travail. Dans chacune des versions, qui sont ci-après appelées la version de Cobb-Douglas (C-D) et la version de la productivité du travail (PT), on ne présume pas à priori que les élasticités du capital et du travail sont égales pour toutes les industries. Par conséquent, les élasticités sont également indexées pour les 88 industries.

Bien que les analyses fondées sur des données de panel posent souvent comme hypothèse que les coefficients de régression sont égaux pour toutes les unités, cette hypothèse doit être vérifiée lorsque l'ensemble de données KLEMS est utilisé. Dans de nombreuses applications, comme les analyses utilisant la *Panel Study on Income Dynamics* (PSID), ou des panels établis pour des provinces ou des États, l'hypothèse est probablement une bonne approximation. Dans ces études, les unités, qui sont soit des personnes soit des régions économiques, ont souvent beaucoup en

commun, ce qui rend plausible l'hypothèse que leurs paramètres de pente dans les équations de régression sont égaux[7].

Dans le cas des données KLEMS, il convient d'analyser davantage l'hypothèse parce que selon la théorie économique, il est probable que les unités (industries) présentent des différences quant aux élasticités des services du capital et des services du travail. Notamment, les industries à prédominance de capital, comme la fabrication ou l'extraction de pétrole et de gaz, ont probablement une élasticité du capital plus élevée que celle des industries à prédominance de travail, comme les finances, les assurances et les services immobiliers. Étant donné ces différences, l'hypothèse que les élasticités devraient être égales d'une industrie à l'autre semble trop restrictive.

Néanmoins, l'hypothèse que les élasticités pourraient être égales est examinée dans un modèle à effets fixes. Généralement, un test F pour la contrainte est utilisé pour vérifier statistiquement l'hypothèse. Cette approche présente toutefois des difficultés lorsque l'ensemble de données KLEMS est utilisé. Premièrement, en estimant la relation au moyen d'un modèle à effets fixes traditionnel, on présume que la variance d'erreur est égale dans toutes les industries. D'après les figures 4 à 6, il semble s'agir d'une hypothèse audacieuse, qui ne s'applique probablement pas aux données. Deuxièmement, la présence d'observations atypiques se traduira par des résidus non normaux qui biaiseront les résultats du test. Un estimateur par les moindres carrés pondérés (MCR) est utilisé pour atténuer cet effet; toutefois, les différences dans l'ampleur des variances d'une industrie à l'autre continueront de poser problème[8].

Comme l'hypothèse que les élasticités du capital et du travail sont égales dans toutes les industries semble trop restrictive, des estimations sont d'abord produites pour chaque industrie individuellement, ce que signifie que 88 régressions sont effectuées. Ensuite, on utilise les distributions des estimations paramétriques et le test F par les MCR pour revoir l'hypothèse que les paramètres sont égaux dans toutes les industries.

---

7. Pour être précis, l'hypothèse est que même s'il peut y avoir une certaine variation entre les unités, cette variation n'est pas assez grande pour rejeter l'hypothèse nulle que les coefficients sont égaux pour l'ensemble des unités.
8. La méthode des moindres carrés pondérés réduit l'incidence des points aberrants et des points leviers en modifiant les coefficients de pondération dans la régression par les MCO (moindres carrés ordinaires). Elle minimise $\min \sum_{t=1}^{T} w_t e_t^2$ où w_t est choisi de telle sorte que les observations atypiques ont un coefficient de pondération faible, alors que les « bonnes » données ont un coefficient de pondération élevé. Il convient de noter que la méthode traditionnelle des moindres carrés ordinaires (MCO) suppose que $w_t = 1 \quad \forall t$. Dans le présent document, le système de pondération employé est simple : $w_t = \begin{cases} = 1 & Non - Aberrant \\ = 0 & Aberrant \end{cases}$. Des systèmes plus complexes peuvent être utilisés au besoin. Pour plus de renseignements, voir Chen, 2002; Rousseeuw et Leroy, 1987.

6. *Estimation en présence d'observations atypiques*

L'analyse préliminaire indique que l'ensemble de données contient des observations atypiques. La présence de telles observations aura des répercussions importantes sur la façon d'utiliser les données KLEMS (capital, travail, énergie, matières et services) pour établir des estimations de la productivité totale des facteurs. Notamment, les observations atypiques peuvent influer sur les estimations paramétriques et l'inférence lorsque la méthode des moindres carrés ordinaires (MCO) est utilisée (voir Sapra, 2003; Tsay, Pena et Pankratz, 2000; Chen et Liu, 1993; Pena, 1990; Rousseeuw et Leroy, 1987; Tasy, 1986; et Rousseeuw, 1984).

Pour contrer cette influence, on utilise dans le présent document un estimateur à point de rupture élevé. Le point de rupture d'un estimateur est une mesure de la sensibilité de ce dernier aux observations atypiques. Il représente la proportion de l'échantillon employé pour l'estimation qui peut être atypique avant que l'estimateur ne soit touché.

6.1 Estimateur MCO

L'estimateur MCO n'est pas un estimateur à point de rupture élevé. La fonction économique pour l'estimation par les MCO minimise la somme résiduelle des carrés :

$$\min \sum_{t=1}^{T} e_t^2.$$

Comme les observations atypiques peuvent se traduire par des erreurs importantes, le problème de la minimisation oriente nécessairement la régression linéaire vers une seule observation qui est arbitrairement grande. Par conséquent, en incluant une seule observation atypique, il est possible de modifier les estimations paramétriques de la pente et de l'ordonnée à l'origine[9]. Comme un unique point de données peut influer sur l'estimateur MCO, son point de rupture est de 0 % (Rousseeuw et Leroy, 1987).

Parce que le point de rupture de l'estimateur MCO est 0, les estimations ne peuvent être utilisées à des fins d'inférence ou d'identification des observations atypiques s'il y en a. Fait important à noter pour les besoins du présent document, des estimations paramétriques biaisées peuvent occulter des observations atypiques dans des statistiques diagnostiques destinées à les révéler. Notamment, pour les observations atypiques assez grandes, l'estimateur MCO est amené vers le point de donnée problématique, de sorte que les statistiques diagnostiques résiduelles fondées sur une régression par les MCO peuvent indiquer que l'observation atypique est un bon point de donnée et que les « bonnes données » posent problème. En outre, lorsqu'il y a plus d'une observation atypique, ou un groupe d'observations atypiques, elles peuvent interagir de façons complexes, en augmentant ou en annulant l'influence exercée par chacune d'elles. Ainsi, un point aberrant peut en cacher un autre, ce qu'on appelle l'effet de masque. De même, un groupe d'observations atypiques peut orienter la régression vers lui, ce qu'on appelle l'effet d'entraînement. En pareil cas, les estimations MCO font qu'il est difficile de déterminer quels

9. Voir Rousseeuw et Leroy, 1987.

points de données sont « bons » et lesquels sont atypiques (Hadi, 1992; Rousseeuw et van Zomeren, 1990; et Rousseeuw et Leroy, 1987).

6.2 Estimateur S

Pour surmonter ces problèmes, il faut un estimateur qui résiste aux observations atypiques. On peut alors choisir parmi plusieurs estimateurs : estimateurs par les moindres carrés médians, estimateurs M, estimateurs par les moindres carrés tronqués et estimateurs S. Ces estimateurs utilisent des fonctions des données qui, de diverses façons, réduisent l'incidence des observations atypiques.

Ils sont souvent qualifiés d'estimateurs robustes parce qu'ils sont insensibles aux observations atypiques. Certains d'entre eux, dont les estimateurs S, ont un point de rupture de 50 %, ce qui signifie que jusqu'à la moitié des observations de l'échantillon peuvent être atypiques avant que l'estimateur ne soit touché. Il s'agit d'une grande amélioration par rapport à l'estimateur MCO, qui est sensible à la présence d'une seule observation atypique[10].

Dans le présent document, l'estimateur S de Rousseeuw et Yohai (1984) est utilisé. Il s'agit d'un estimateur à point de rupture élevé qui sélectionne les estimations paramétriques dans le but de minimiser la dispersion des résidus. Fait important, cet estimateur est robuste aux observations atypiques dans la variable dépendante et dans les variables indépendantes.

L'estimateur S minimise la dispersion des résidus de régression de sorte que :

$$\hat{\theta}_S = \arg\min_{\theta} S(\theta)$$

où $S(\hat{\theta})$ est une estimation de la dispersion de la distribution.

L'estimation S est sélectionnée en appliquant à l'échantillon de données un algorithme qui tire un sous-échantillon de données pour calculer les coefficients de régression. Ces coefficients sont utilisés pour résoudre de façon itérative une équation supplémentaire et produire une estimation de $S(\hat{\theta})$. Lorsque l'algorithme est appliqué de façon répétée, des vecteurs des estimations de dispersion sont créés. Comme les observations atypiques augmentent nécessairement la dispersion des résidus, un ensemble robuste d'estimations paramétriques peut être trouvé en sélectionnant les coefficients de régression qui correspondent à la plus petite valeur $S(\hat{\theta})$[11]. L'algorithme fournit des estimations paramétriques robustes qui sont asymptotiquement équivalentes aux estimations établies à partir d'un modèle de régression gaussien normal.

---

10. Il convient de prendre note que l'estimateur par les moindres écarts absolus (MEA) peut avoir le même point de rupture que l'estimateur MCO (moindres carrés ordinaires). Bien que l'estimateur MEA fournisse la médiane dans un échantillon univarié et qu'il puisse être robuste aux observations atypiques dans la variable dépendante, pour une observation atypique assez grande dans la variable indépendante, la pente de l'estimateur MEA dépassera le point problématique. Pour plus de renseignements, voir Rousseeuw et Leroy, 1987.
11. Pour obtenir des renseignements techniques sur l'algorithme employé pour calculer les estimations S utilisées dans le présent document, voir la section « Proc RobustReg » dans le manuel de l'utilisateur de SAS9.

6.3 *Repérage des points aberrants et des points leviers*

Puisque les observations atypiques peuvent n'être présentes que dans un sous-ensemble des variables utilisées pour l'analyse, une distinction est établie entre les observations atypiques présentes dans les variables indépendantes (points leviers) et celles présentes dans la variable de réponse (points aberrants). Comme on l'a vu précédemment, il est nécessaire d'utiliser des estimateurs robustes pour repérer les points leviers et les points aberrants (Hadi, 1994, 1992; Rousseeuw et van Driesen, 1999).

Pour repérer les points leviers, on mesure la distance robuste d'une observation par rapport à un emplacement central robuste. La statistique de distance robuste utilisée ici est la suivante :

$$RD(x_i) = \left[\left(x_i - T(x_i)\right)' C(x_i)^{-1} \left(x_i - T(x_i)\right) \right]^{0.5}$$

où $T(x_i)$ et $C(x_i)$ sont des estimations multivariées robustes de l'emplacement central et de la matrice de dispersion respectivement[12]. Les points leviers sont alors définis comme étant les distances robustes qui sont supérieures à la valeur critique correspondante, laquelle est fondée sur la loi du chi-carré :

$$Crit(p) = \left(\chi^2_{p,1-\alpha}\right)^{0.5}$$

où p est le nombre de variables explicatives et α est le seuil de signification préétabli. Lorsque $RD(x_i)$ est supérieur à $Crit(p)$, un point levier est présumé être présent. Bien que le choix du seuil de signification soit arbitraire, dans le présent document tous les tests concernant les points leviers ont un seuil de signification de 5%.

La détection des points aberrants est fondée sur les résidus d'une régression et sera donc influencée par la forme fonctionnelle, les contraintes paramétriques et les variables explicatives qui sont incluses. Les estimations S sont utilisées pour créer un vecteur résiduel $\hat{\varepsilon}$. Pour chaque $\hat{\varepsilon}_t$ en $\hat{\varepsilon}$, la valeur absolue du résidu est comparée au double de l'estimation du paramètre d'échelle $S(\hat{\theta}_S)$. Lorsque la valeur absolue de $\hat{\varepsilon}_t$ est supérieure à ce seuil, l'observation est considérée comme un point aberrant.

12. Dérivations et notation selon Chen (2002). Les estimations de $RD(x_i)$ sont établies au moyen de SAS9, qui utilise l'algorithme de Rousseeuw et Van Dreissen (1999) pour calculer des statistiques multivariées robustes. Pour plus de renseignements, voir Chen (2002).

7. *Estimations paramétriques de la PTF*

L'examen des estimations paramétriques établies à partir de l'équation de Cobb-Douglas (C-D) et de l'équation de productivité du travail (PT) dans la présente section montre dans quelle mesure les estimations MCO (moindres carrés ordinaires) tirées de la base de données KLEMS (capital, travail, énergie, matières et services) sont biaisées en raison des points aberrants et des points leviers. La comparaison débute par l'examen des industries échantillonnées, suivi d'une analyse du panel complet du secteur des entreprises. Aux fins de cette analyse, seulement 88 des 89 industries incluses dans la base de données KLEMS sont utilisées. La dernière industrie, soit le secteur non commercial, est exclue. Elle est toutefois réintégrée dans le panel dans les sous-sections subséquentes où les observations atypiques sont analysées en fonction de leur nombre et des périodes où elles sont présentes.

L'analyse porte sur la façon dont les observations atypiques influent sur les estimations MCO et peuvent être révélées par l'application de techniques d'estimation différentes. Elle ne vise pas à déterminer si l'équation est spécifiée correctement ou non. Notamment, la spécification peut ne pas tenir compte de la dynamique de la productivité totale des facteurs (PTF), de l'utilisation du capital au cours du cycle économique ou de l'incidence des économies d'échelle. Ces aspects importants, qui peuvent occasionner un biais de variables omises, ne sont pas visés par l'examen économétrique, lequel consiste plutôt à illustrer les répercussions que les observations atypiques peuvent avoir sur les estimations paramétriques si elles ne sont pas prises en considération. La spécification de la PTF employée ici est couramment utilisée dans les travaux de recherche et est concordante avec les spécifications non paramétriques de la PTF.

Avant l'examen du panel d'industries, les industries échantillonnées sont utilisées pour illustrer, *ceteris paribus*, l'incidence que peut avoir l'application d'une nouvelle stratégie d'estimation. L'ensemble du panel du secteur des entreprises est examiné ensuite, puis un certain nombre de conclusions générales sont tirées au sujet des sources d'observations atypiques.

7.1 Estimations de la PTF pour les industries échantillonnées

Pour l'industrie de la fabrication de textiles et de produits textiles, l'estimation MCO de *ptf* selon l'équation de C-D se chiffre à 2,43 %, soit 0,43 point de pourcentage plus élevée que l'estimation S qui s'établit à 2,0 % (tableau 5). Il y a un écart entre les estimations des élasticités, qui est toutefois plus faible que l'écart entre les estimations de *ptf*.

La méthode des MCO entraîne une surestimation de *ptf* pour l'industrie de la fabrication de textiles et de produits textiles, mais une sous-estimation pour les autres industries échantillonnées. Pour l'industrie de la fabrication de matériel informatique et périphérique, l'estimation MCO de *ptf* est de 8,88 % alors que l'estimation S est de 13,17 %, soit un écart de 4,30 points de pourcentage. Il s'agit de l'écart le plus important observé dans le secteur des entreprises. En outre, l'estimation S de l'élasticité du travail est de 0,46, ce qui est inférieur de 0,63 point de pourcentage à l'estimation MCO qui se chiffre à 1,09. Qui plus est, alors que l'hypothèse de production à rendements constants à l'échelle ne semble pas bien s'ajuster aux données lorsque la méthode des MCO est employée, l'estimation S semble supporter la contrainte.

Tableau 5
Estimations paramétriques selon l'équation de Cobb-Douglas

Fabrication de textiles et de produits textiles	MCO[1]	Estimateur S
ptf	2,43	2,00
$\hat{\beta}_L$	1,11	1,05
$\hat{\beta}_K$	-0,04	-0,01
Fabrication de matériel informatique et périphérique	**MCO[1]**	**Estimateur S**
ptf	8,88	13,17
$\hat{\beta}_L$	1,09	0,46
$\hat{\beta}_K$	0,59	0,55
Fabrication de matériel ferroviaire roulant	**MCO[1]**	**Estimateur S**
ptf	-1,12	1,06
$\hat{\beta}_L$	1,19	1,09
$\hat{\beta}_K$	0,22	0,12

1. Moindres carrés ordinaires.
Source : Statistique Canada.

Pour l'industrie de la fabrication de matériel ferroviaire roulant, l'écart est moins prononcé. Toutefois, l'estimation de *ptf* change de signe lorsque les observations atypiques sont prises en considération. L'estimation MCO de *ptf* dans cette industrie est de -1,12 % par année. Lorsqu'on tient compte des observations atypiques dans les données, l'estimation S de *ptf* grimpe de 2,18 points de pourcentage pour atteindre 1,06 %. Les estimations S pour l'élasticité des services du capital et des services du travail sont inférieures aux estimations MCO.

Pour l'industrie de la fabrication de textiles et de produits textiles, les estimations de croissance de *ptf* selon l'équation de PT sont similaires aux estimations selon l'équation de C-D (tableau 6). L'estimation MCO montre que la croissance de *ptf* est de 2,48 %, soit 0,42 point de pourcentage plus élevée que selon l'estimation S qui se chiffre à 2,06. L'estimation de l'élasticité des services du capital est près de zéro selon les deux méthodes d'estimation, mais l'estimation MCO est presque trois fois plus élevée que l'estimation S.

Tableau 6
Estimations selon l'équation de productivité du travail

Fabrication de textiles et de produits textiles	MCO[1]	Estimateur S
ptf	2,48	2,06
$\hat{\beta}_K$	-0,08	-0,03
Fabrication de matériel informatique et périphérique	**MCO[1]**	**Estimateur S**
ptf	14,15	13,26
$\hat{\beta}_K$	0,06	0,54
Fabrication de matériel ferroviaire roulant	**MCO[1]**	**Estimateur S**
ptf	-0,47	1,27
$\hat{\beta}_K$	-0,14	-0,04

1. Moindres carrés ordinaires.
Source : Statistique Canada.

Alors que selon l'équation de C-D la méthode des MCO sous-estime la croissance de *tfp* pour l'industrie de la fabrication de matériel informatique et périphérique, l'estimation de croissance de *ptf* selon l'équation de PT est de 14,15 %, soit 0,89 point de pourcentage de plus que l'estimation S qui s'établit à 13,26 %. Comme dans le cas des estimations selon l'équation de C-D, l'estimation MCO de l'élasticité du capital (0,06) est inférieure à l'estimation S (0,54).

Pour l'industrie de la fabrication de matériel ferroviaire roulant, l'estimation MCO de la croissance de *ptf* selon l'équation de PT est de -0,47 %; elle est négative comme l'estimation de type C-D, mais elle a une amplitude moitié moins élevée. L'estimation S augmente de 1,74 point de pourcentage pour atteindre 1,27 %, ce qui est similaire à l'estimation sans contrainte de type C-D. L'estimation de l'élasticité du capital passe de -0,14 à 0,04 lorsque l'estimateur S est employé.

Les industries échantillonnées montrent l'incidence que les points aberrants et les points leviers peuvent avoir sur les estimations paramétriques lorsque la base de données KLEMS est utilisée. Les observations atypiques influeront sur le processus d'estimation par les MCO à partir de ces données. Bien que, dans l'industrie de la fabrication de matériel informatique et périphérique, l'estimation de croissance de *ptf* semble augmenter lorsque la contrainte des rendements constants à l'échelle est imposée sur la fonction de production, cela ne devrait pas être généralisé. L'estimation MCO de type PT de l'élasticité du capital pour cette industrie est inférieure à l'estimation S, et les estimations de *ptf* pour les autres industries échantillonnées continuent d'être influencées par les observations atypiques.

7.2 Estimations de la PTF pour le panel d'industries

Les estimations paramétriques établies au moyen de l'estimateur MCO et de l'estimateur S pour les industries échantillonnées suffisent à démontrer que les observations atypiques peuvent avoir des répercussions importantes sur une industrie donnée. Toutefois, les résultats obtenus pour les industries échantillonnées ne fournissent pas assez d'information pour formuler des généralisations quant à l'incidence des points aberrants et des points leviers sur le panel des industries du secteur des entreprises.

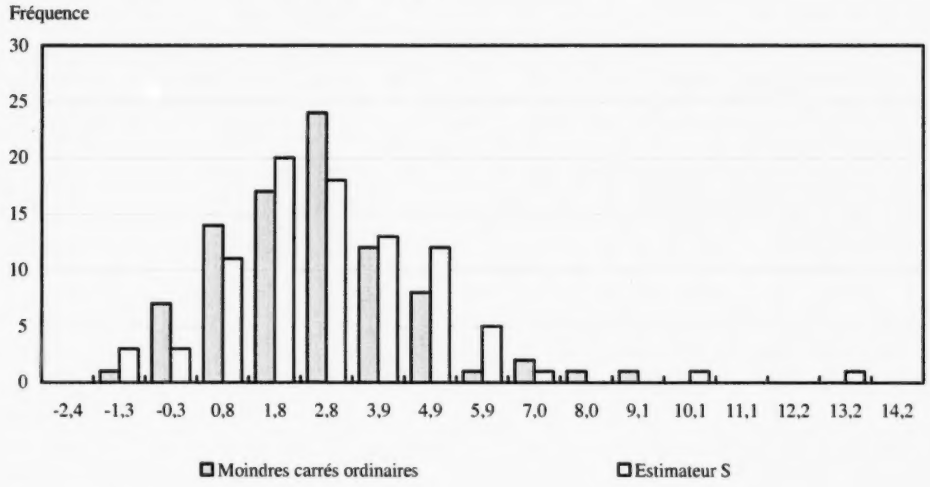
Il est difficile d'analyser le secteur des entreprises étant donné, comme on l'a mentionné auparavant, que l'intensité de capital varie selon les industries, ce qui laisse supposer que l'élasticité du capital est également différente d'une industrie à l'autre. Par conséquent, les données ne sont pas regroupées initialement, et des régressions sont effectuées pour chaque industrie séparément. Les estimations établies au moyen de l'estimateur MCO et de l'estimateur S sont regroupées, et leurs distributions examinées. L'analyse porte en majeure partie sur ces distributions parce qu'il est difficile de procéder à des tests basés sur des résidus en raison de la présence d'observations atypiques dans l'ensemble de données. Malgré tout, on utilise la méthode des moindres carrés pondérés (MCR) pour tester statistiquement la validité des contraintes inter-équations. Cependant, la méthode des MCR employée ici pose nécessairement la contrainte que les variances des industries soient égales, contrainte qui est probablement trop forte.

L'analyse est axée sur les distributions des estimations paramétriques inter-industries parce qu'elles permettent de tester officieusement un certain nombre d'hypothèses. Premièrement, si les points aberrants et les points leviers ont tendance à toucher toutes les industries de façon similaire, le point central des distributions de l'estimateur MCO et de l'estimateur S sera différent. Deuxièmement, si l'estimateur MCO tend à omettre les estimations paramétriques élevées, ce qui est le cas pour les estimations de croissance de *ptf* selon l'équation de C-D, les queues de la distribution des estimations S seront plus épaisses que celles des estimations MCO. Troisièmement, si l'hypothèse que la croissance de *ptf*, ou les élasticités des services du capital et du travail, est égale d'une industrie à l'autre est valide, les distributions devraient présenter une faible variance. Elles devraient être concentrées autour de leur point central[13].

Selon l'équation de C-D, les estimations de la croissance de *tfp* semble être influencées négativement par les observations atypiques lorsque l'estimateur MCO est utilisé, lequel fournit un plus grand nombre de faibles estimations que l'estimateur S (tableau 9). De plus, l'estimateur MCO ne semble pas estimer adéquatement la croissance de *ptf* dans les industries où celle-ci est la plus forte. Le taux de croissance moyen pour les estimations MCO est de 2,08 %, contre 2,43 % pour les estimations S.

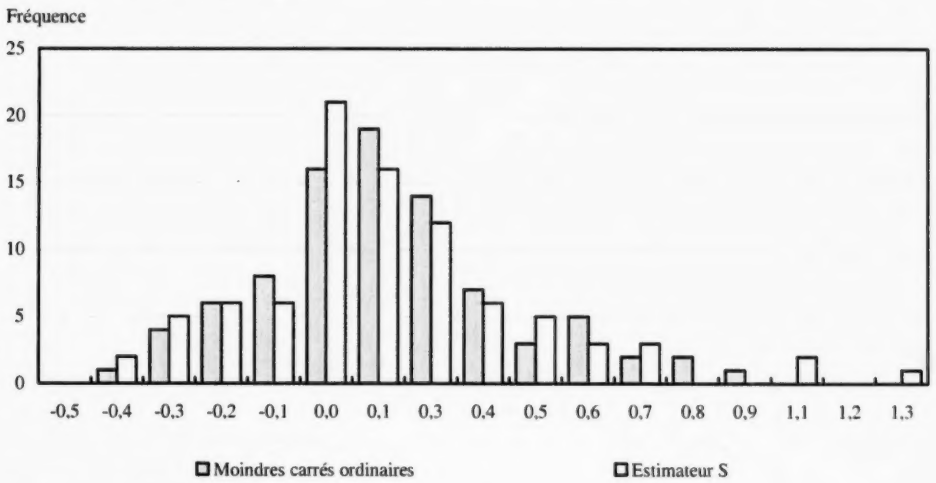
13. Le tableau complet des estimations paramétriques figure à l'annexe A.

Figure 9
Distribution des estimations de la productivité totale des facteurs selon la fonction de production Cobb-Douglas



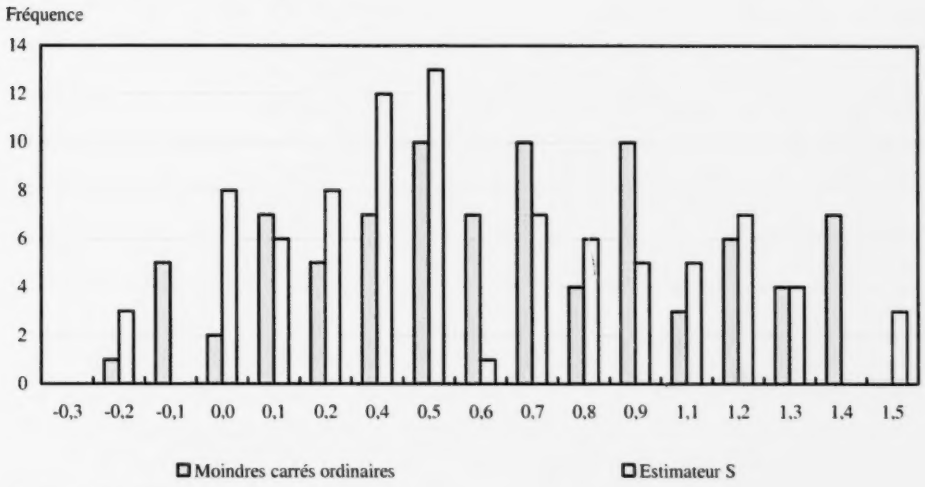
Source : Statistique Canada.

Figure 10
Distribution des estimations de l'élasticité du capital selon la fonction de production Cobb-Douglas



Source : Statistique Canada.

Figure 11
Distribution des estimations de l'élasticité du travail selon la fonction de production Cobb-Douglas



Source : Statistique Canada.

Toutefois, comme l'estimateur MCO semble sous-estimer les taux de croissance de *ptf* les plus élevés, la comparaison des moyennes des taux de croissance de *ptf* selon l'estimateur S et l'estimateur MCO peut refléter l'incidence d'un petit nombre d'observations positives de grande taille. Si on tronque les cinq taux de croissance positifs les plus élevés et si on calcule de nouveau la moyenne des taux de croissance de *ptf* , un portrait cohérent se dessine : la moyenne des taux de croissance de *ptf* se chiffre à 1,80 avec l'estimateur MCO et à 2,10 avec l'estimateur S. Dans chaque cas, l'estimateur MCO sous-estime la croissance moyenne de *ptf* .

Si l'estimateur MCO est appliqué à la base de données KLEMS, les estimations du taux de croissance de *ptf* ont tendance à sous-estimer la croissance de *ptf* et ne signalent pas les industries où la croissance de *ptf* est la plus forte. En outre, les écarts ne semblent pas être aléatoires puisque les estimations MCO sont, en moyenne, inférieures aux estimations S.

Dans le cas des estimations de l'élasticité du capital selon l'équation de C-D, il semble que l'estimateur MCO puisse sous-estimer l'élasticité tout en ne parvenant pas à reconnaître les valeurs les plus grandes (figure 10). Cependant, les moyennes des estimations MCO et des estimations S sont similaires, soit 0,10 et 0,11 respectivement. Le calcul d'une moyenne tronquée excluant les cinq estimations les plus élevées n'y change rien.

La similitude entre les estimations MCO et les estimations S est inattendue étant donné la présence de points leviers et de points aberrants dans l'ensemble de données. Cette contradiction apparente peut être expliquée par l'élasticité à faible moyenne qui est présumée. La moyenne proche de zéro peut résulter de faibles écarts d'une année à l'autre. Cependant, il est également possible que, même si les estimations MCO et les estimations S sont différentes dans un grand nombre d'industries, cette différence soit aléatoire et non systématique.

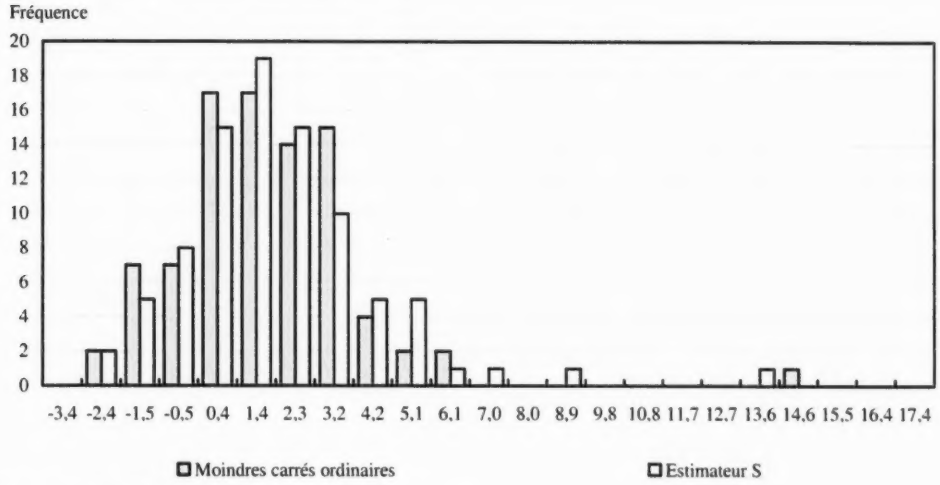
Contrairement aux estimations de la croissance de *ptf* et aux estimations de l'élasticité du capital, les estimations de l'élasticité du travail selon l'équation de C-D suggèrent que l'estimateur MCO surestime l'élasticité moyenne du travail (figure 11). L'estimation MCO moyenne est de 0,59, soit 0,07 point de pourcentage de plus que l'estimation S moyenne qui se chiffre à 0,52. Le même écart est observable lorsque les cinq estimations les plus élevées sont tronquées dans l'échantillon.

Les estimations de croissance de *ptf* selon l'équation de PT présentent un profil légèrement différent de celui des estimations selon l'équation de C-D. Avec la contrainte des rendements constants à l'échelle, les distributions se centralisent, bien que l'estimateur MCO continue de sous-estimer la croissance de *ptf* par rapport à l'estimateur S. La moyenne des estimations MCO de la croissance de *ptf* est de 1,23 %, soit 0,15 point de pourcentage de plus que la moyenne des estimations S qui se chiffre à 1,38 %. Comme dans le cas des estimations selon l'équation de C-D, il y a un petit nombre d'estimations de croissance de *ptf* qui sont élevées. Lorsqu'on supprime les cinq estimations de croissance de *ptf* les plus élevées de l'estimateur MCO et de l'estimateur S, les moyennes tombent à 0,89 % et 1,0 % respectivement. Bien que les estimations soient moins dispersées lorsque la contrainte des rendements constants à l'échelle est imposée, elles restent sensiblement différentes d'une industrie à l'autre et continuent de suggérer la présence d'un écart systématique.

Les estimations de l'élasticité du capital établies avec l'estimateur MCO et l'estimateur S à partir de l'équation de PT sont sensiblement différentes d'une industrie à l'autre. Malgré cela, les moyennes sont similaires (estimations MCO : 0,28; estimations S : 0,31) et se rapprochent de la part du capital dans le revenu du secteur des entreprises. De plus, l'élasticité du travail présumée se chiffre à 0,72 (estimations MCO) ou à 0,69 (estimations S), valeur proche de la valeur de la part du travail dans le revenu du secteur des entreprises. Comme les estimations selon l'équation de C-D, les estimations de l'élasticité selon l'équation de PT ne laissent pas présumer la présence d'un biais important dans les estimations MCO des différentes industries.

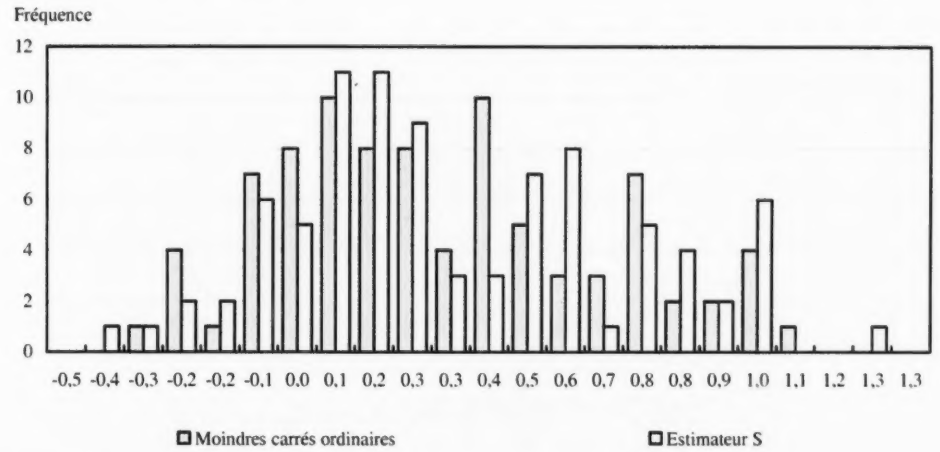
L'analyse des distributions des paramètres selon l'équation de C-D et l'équation de PT montre que l'estimateur MCO tend à sous-estimer la croissance de *ptf* par rapport à l'estimateur S. Il y a sous-estimation dans la moyenne des estimations et dans les estimations paramétriques élevées pour certaines industries. Cette tendance suggère que les innovations défavorables de grande envergure, qui ne représentent pas la majorité des données, peuvent avoir des répercussions importantes sur les estimations paramétriques de la croissance de *ptf*. Notamment, les estimations MCO sont établies à partir des moyennes des variables indépendantes et dépendantes. Si des événements défavorables ponctuels et rares, comme des récessions, influent sur ces moyennes, alors les estimations MCO de la croissance *ptf* seront également biaisées. En outre, si certaines industries sont plus sensibles aux chocs économiques occasionnels, alors l'écart entre leurs estimations de croissance de *ptf* devrait être plus grand.

Figure 12
Distribution des estimations de la productivité totale des facteurs selon la fonction de productivité du travail



Source : Statistique Canada.

Figure 13
Distribution des estimations de l'élasticité du capital selon la fonction de productivité du travail



Source : Statistique Canada.

Les distributions suggèrent que les estimations MCO, puisqu'elles utilisent toutes les données disponibles, sont influencées par les périodes d'activité économique irrégulière. Cependant, l'estimateur S est axé sur les observations qui représentent la majorité des données et n'est pas touché par le même problème. Cette différence fait ressortir un problème important touchant l'estimation de *ptf* .

En période de récession, les fonctions de réponse des agents économiques peuvent être différentes de celles observables en période d'expansion. Par conséquent, si un chercheur essaie de saisir la croissance à long terme à partir de hausses touchant des facteurs intangibles ou difficiles à mesurer, il peut être mal avisé d'inclure des périodes de récession dans l'ensemble de données utilisé pour l'estimation. Bien qu'un examen complet de cette question n'entre pas dans le champ d'étude du présent document, les sections subséquentes montreront que les points aberrants et les points leviers tendent à se regrouper autour de chocs agrégés, ce qui laisse entendre que ces périodes constituent une source importante d'observations inhabituelles.

Les distributions des estimations paramétriques établies selon les deux équations suggèrent que la croissance de *ptf*, de même que les élasticités du capital et du travail, diffèrent d'une industrie à l'autre. Les estimations ont tendance à être dispersées au lieu d'être regroupées autour d'un emplacement central. En présence de ces types d'écarts entre les estimations paramétriques, la validité des contraintes inter-équations est incertaine. Néanmoins, l'hypothèse est testée officiellement pour la croissance de *ptf* et les élasticités selon les deux équations.

L'équation de C-D rejette l'hypothèse que la croissance de *ptf* et l'élasticité du travail sont égales d'une industrie à l'autre (tableau 7). Cependant, le test F ne rejette pas l'hypothèse pour l'élasticité du capital, ce qui ne concorde pas avec la dispersion des estimations paramétriques tirées des régressions individuelles. Pour l'équation de PT, l'hypothèse que la croissance de *ptf* et l'élasticité du capital sont égales est rejetée au seuil de 5 %.

À l'exception de l'élasticité du capital dans l'équation de C-D, les tests F soutiennent la notion que les estimations paramétriques sont différentes d'une industrie à l'autre. Le non-rejet de l'hypothèse nulle pour l'élasticité du capital dans l'équation de C-D est surprenant étant donné la dispersion des estimations individuelles. Il peut découler du fait que la technique des MCR réduit la variabilité du changement dans le stock de capital, lequel a par avance tendance à afficher une faible variation, ce qui fait que les estimations paramétriques ne sont pas statistiquement significatives. Pour la majorité des estimations paramétriques de l'élasticité du capital, c'est peut-être le cas; il reste toutefois une minorité d'estimations qui sont significatives au seuil de 5 %.

Néanmoins, la dispersion des estimations paramétriques et des résultats des tests F indique avec force que les données ne devraient pas être regroupées et que les paramètres ne devraient pas être affectés de contraintes inter-équations. Pour décider s'ils doivent traiter les données en panel ou les utiliser dans des régressions individuelles, les chercheurs devront déterminer s'ils peuvent raisonnablement imposer la contrainte que les variances des équations individuelles sont égales.

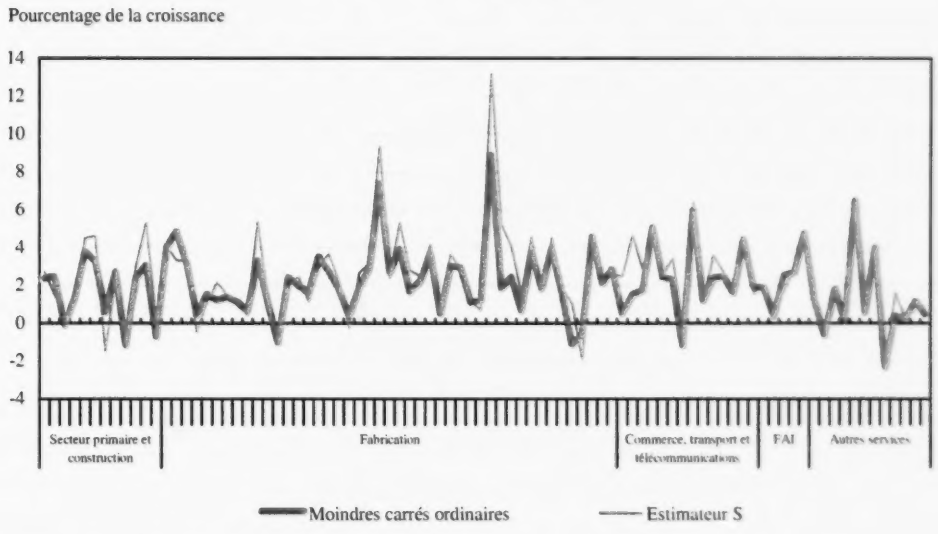
Tableau 7
Tests F des contraintes inter-équations moindres carrés pondérés (MCR)

Hypothèse	Test F
Équation de Cobb-Douglas	
$ptf_i = ptf$	1,60 (0,00)
$\hat{\beta}_{K,i} = \hat{\beta}_K$	1,03 (0,41)
$\hat{\beta}_{L,i} = \hat{\beta}_L$	2,04 (0,00)
Équation de productivité du travail	
$ptf_i = ptf$	1,84 (0,00)
$\hat{\beta}_{K,i} = \hat{\beta}_K$	2,72 (0,00)

Nota : Valeurs P entre parenthèses.
Source : Statistique Canada.

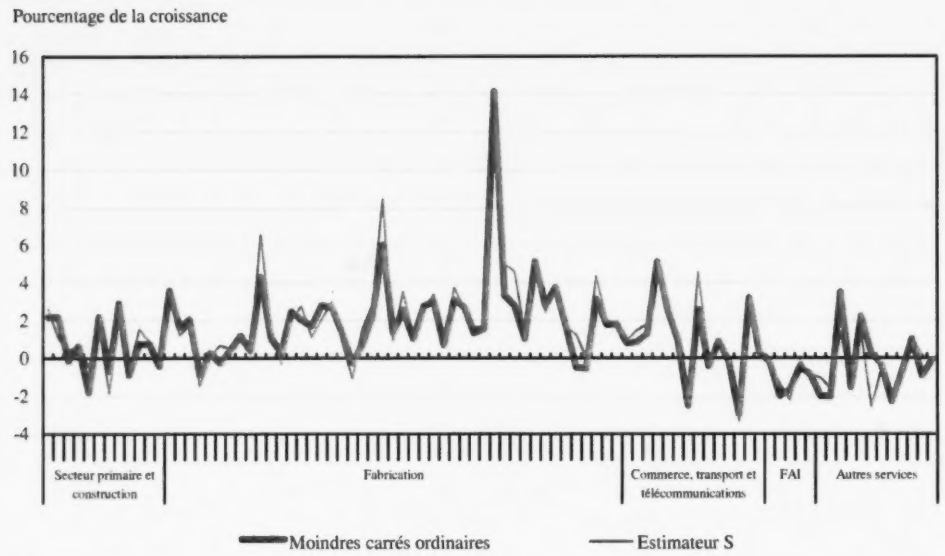
Il convient de souligner que la tendance des estimations paramétriques du taux de croissance de *ptf* diffère d'une industrie à l'autre selon que la contrainte des rendements constants à l'échelle est imposée ou non (figures 14 et 15). Alors que les deux équations montrent que la croissance de *ptf* dans le secteur de la fabrication tend à être plus forte que dans les autres secteurs, lorsque la contrainte des rendements constants à l'échelle est imposée sur la fonction de production, les estimations du secteur primaire et du secteur des services deviennent faibles ou négatives.

Figure 14
Estimations de la productivité totale des facteurs selon la fonction de production de l'équation de Cobb-Douglas par industrie



Source : Statistique Canada.

Figure 15
Estimations de la productivité totale des facteurs selon la fonction de production de l'équation de productivité du travail par industrie



Source : Statistique Canada.

Par conséquent, la contribution des différents secteurs à la productivité du secteur des entreprises dépend de la validité de la contrainte des rendements constants à l'échelle. Si cette contrainte est valide, alors la productivité du secteur des entreprises est dictée presque complètement par le secteur de la fabrication et tend à être retenue par les autres secteurs. Toutefois, si cette supposition est invalide, alors ce sont seulement quelques industries, et non des secteurs, qui diminueront la croissance de la productivité.

Il est important de faire cette distinction parce que Baldwin, Durand et Hosein (2001) ont démontré que les secteurs affichant la plus forte croissance de productivité subissent une baisse de leur poids agrégé avec le temps. Par conséquent, si la contrainte des rendements constants à l'échelle est valide, il peut être nécessaire, afin de maintenir la croissance de la productivité du secteur des entreprises à un niveau donné dans le temps, de connaître une augmentation permanente et accélérée de la croissance de la productivité du secteur de la fabrication.

Il importe également de noter que le groupe d'industries affichant l'écart le plus important entre les estimations S et les estimations MCO du taux de croissance de *ptf* n'est pas le même selon qu'il s'agit de l'équation de C-D ou de l'équation de PT (tableaux 8 et 9). Bien que certaines industries se retrouvent dans les deux groupes, l'amplitude des biais dans les estimations MCO selon l'équation de C-D tend à être plus forte que celle des biais dans les estimations MCO selon l'équation de PT.

Tableau 8
Écarts entre les estimations S et les estimations MCO[1] de la PTF[2] selon l'équation de Cobb-Douglas

Industrie	Écart
Fabrication de matériel informatique et périphérique	4,30
Fabrication de produits électroniques	3,38
Commerce de détail	3,14
Distribution de gaz naturel et réseaux d'aqueduc et d'égout et autres	2,25
Fabrication de matériel ferroviaire roulant	2,18
Extraction de minerais métalliques	-2,05
Vineries	2,03
Commerce de gros	2,00
Fabrication de résines, de caoutchouc synthétique et de fibres et de filaments artificiels et synthétiques	1,98
Fabrication d'appareils ménagers	1,59
Fabrication de sucre et de confiseries	-1,51
Services de gestion des déchets et d'assainissement	-1,39
Fabrication de produits pharmaceutiques et de médicaments	1,39
Extraction de charbon	1,34
Construction de navires et d'embarcations	-1,23
Hébergement et services de restauration	1,23
Services postaux et messageries et services de messagers	1,23

1. Moindres carrés ordinaires.
2. Productivité totale des facteurs.
Source : Statistique Canada.

Tableau 9
Écarts entre les estimations S et les estimations MCO[1] de la PTF[2] selon l'équation de productivité du travail

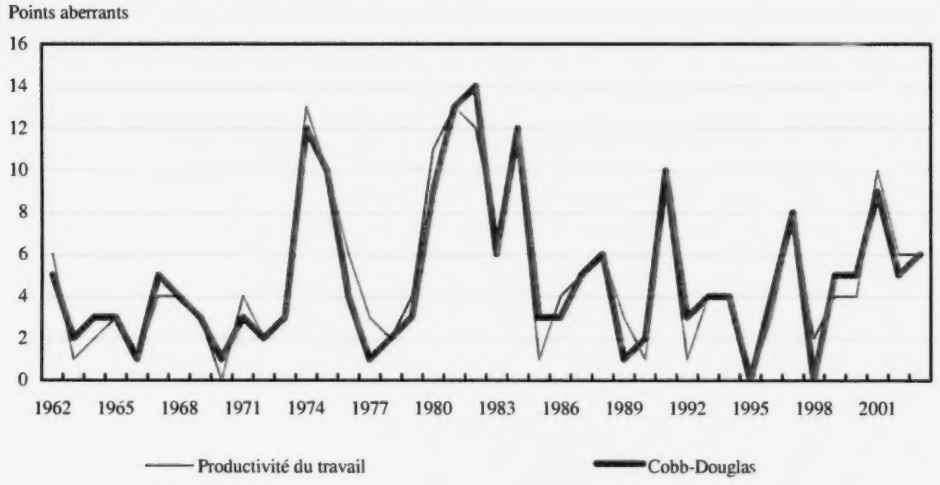
Industrie	Écart
Services d'enseignement, sauf les universités	-2,84
Fabrication de résines, de caoutchouc synthétique et de fibres et de filaments artificiels et synthétiques	2,41
Vineries	2,24
Autres services professionnels, scientifiques et techniques	-2,14
Transport par pipeline	1,99
Fabrication d'appareils ménagers	1,82
Fabrication de matériel ferroviaire roulant	1,74
Fabrication de produits électroniques	1,67
Extraction de pétrole et de gaz	1,33
Fabrication d'autres types de matériel de transport	1,21
Extraction de minerais métalliques	-1,15
Foresterie et exploitation forestière	-1,10
Fondations, groupes de citoyens et organisations professionnelles et similaires	1,07
Fabrication de produits pharmaceutiques et de médicaments	0,96
Services de publicité et services connexes	0,94
Préparation et conditionnement de poissons et de fruits de mer	0,94

1. Moindres carrés ordinaires.
2. Productivité totale des facteurs.
Source : Statistique Canada.

Quoi qu'il en soit, dans les deux cas l'ampleur du biais est non négligeable et suggère que les chocs économiques ne touchent pas toutes les industries également. Fait important, les écarts d'amplitude importants entre l'équation de C-D et l'équation de PT laissent entendre que les estimations MCO ont tendance à sous-estimer la croissance de *ptf* , ce qui appuie la supposition que des chocs défavorables occasionnels, comme des récessions, biaisent de façon importante les estimations MCO de la croissance de *ptf* . En outre, l'incidence de ces chocs varie d'une industrie à l'autre, ce qui laisse supposer que leurs réactions ne sont pas égales. Au contraire, la relation entre les facteurs et la valeur ajoutée semble varier considérablement.

Les distributions et les valeurs des écarts les plus grands suggèrent que des événements occasionnels, comme des récessions, se traduisent par des observations atypiques[14], ce qui laisse présumer que le nombre d'observations atypiques devrait augmenter autour de ces types de chocs économiques occasionnels. C'est effectivement ce qui se produit (figures 16 et 17). On observe une augmentation du nombre de points aberrants dans le PIB réel et la productivité du travail par industrie au moment du premier choc pétrolier au début des années 1970, de la récession de 1980-1981, de la récession de 1990 et vers la fin de la période d'échantillonnage.

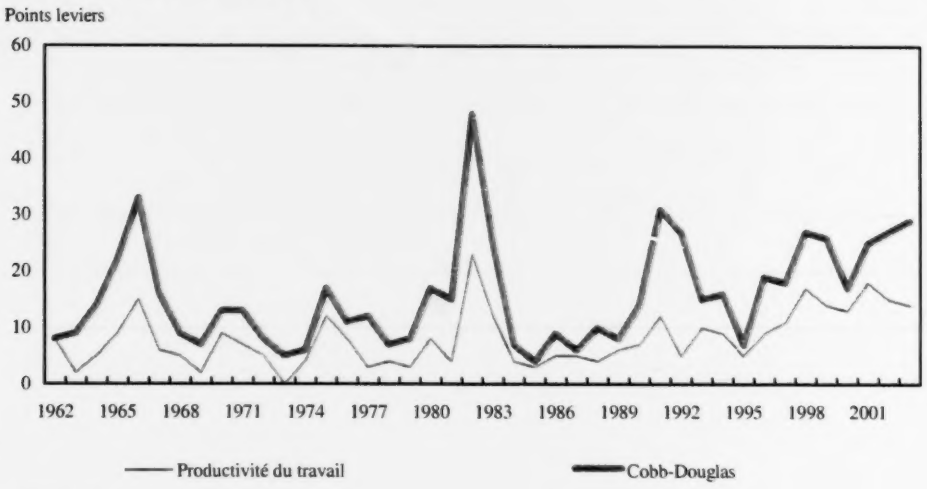
Figure 16
Points aberrants par année



Source : Statistique Canada.

14. Le secteur non commercial est réintégré dans l'ensemble de données afin que le nombre total de points aberrants et de points leviers puisse être examiné. Son inclusion n'a pas d'incidence sur les résultats. Alors, à des fins d'exhaustivité, il est inclus dans le calcul du nombre total de points aberrants et de points leviers trouvés dans l'ensemble de données.

Figure 17
Points leviers par année



Source : Statistique Canada.

Par ailleurs, le nombre de points leviers dans les services du capital, les services du travail et le ratio capital-travail répond également pendant les événements économiques inhabituels[15]. Le nombre de points leviers augmente au milieu des années 1960, au moment des récessions de 1980-1981 et de 1991 et vers la fin de la période d'échantillonnage.

Comme le montrent les figures 16 et 17, les points aberrants et les points leviers sont fréquents. Pour l'équation de C-D et l'équation de PT, il y a 206 points aberrants sur un total de 3 738 observations. Des points aberrants sont donc présents dans 5,5 % des variables de réponse dans l'échantillon. Pour l'équation de C-D, il y a 665 points leviers (17,8 % des observations de l'échantillon), alors que pour l'équation de PT, il y a en 340 (9,1 % des observations de l'échantillon).

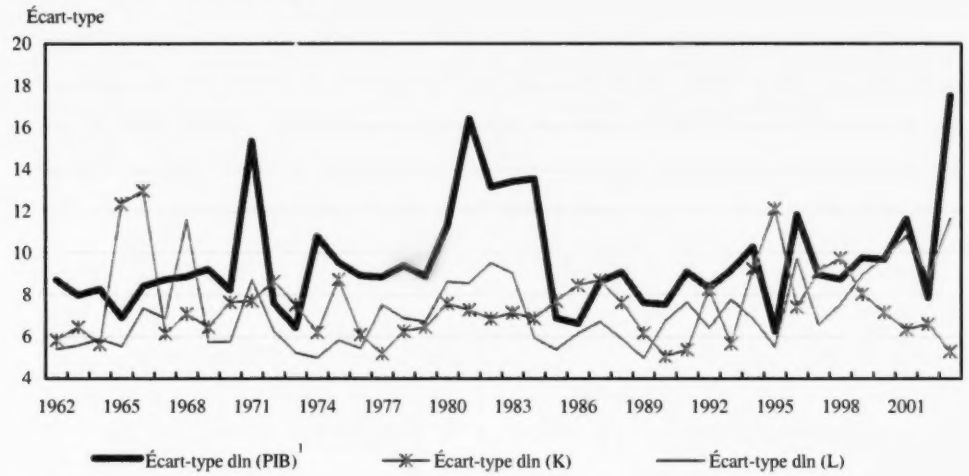
Bien que la proportion de points leviers et de points aberrants dans l'échantillon soit digne de mention, il est important de garder à l'esprit qu'ils n'ont pas à être présents la même année. Même s'il y a des regroupements d'observations atypiques autour d'événements économiques précis, le nombre réel d'observations atypiques dont il faut tenir compte dans l'échantillon est calculé comme le nombre d'années au cours desquelles au moins un type d'observations atypiques (points aberrants ou points leviers) est présent. Lorsque les données sont combinées de cette façon, il y a 506 observations atypiques (13,5 % des observations de l'échantillon) lorsque l'équation de PT est utilisée, et 812 (21,7 % des observations de l'échantillon) lorsque c'est l'équation de C-D. Ces totaux montrent que les points aberrants et les points leviers constituent un élément important de la base de données KLEMS.

15. Une liste présentant les points aberrants et les points leviers par industrie et par année est présentée aux annexes B et C.

L'ampleur des biais les plus importants suggère également que certaines industries sont touchées plus que les autres par les chocs économiques occasionnels. Si c'est le cas, alors la réaction des industries devrait varier davantage lorsque des chocs ponctuels affectent le système. Il est possible de vérifier cette hypothèse de façon informelle en créant une série chronologique des écarts-types pour les différences logarithmiques des industries chaque année. Si certaines industries sont plus sensibles aux chocs économiques, alors l'écart-type devrait augmenter lorsque ces chocs se produisent.

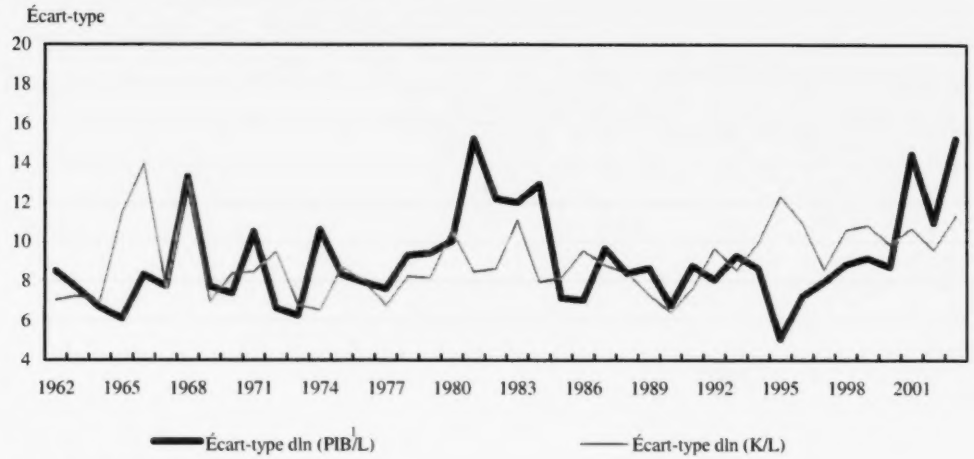
L'examen de la série chronologique fait ressortir la tendance hypothétique (figures 18 et 19). L'effet est plus prononcé pour les variables dépendantes que pour les facteurs. Néanmoins, on observe une augmentation de la dispersion du PIB et des services du travail lors des chocs macroéconomiques du début des années 1970, au début des années 1980 et en 2003. La réponse est toutefois faible lors de la récession de 1991.

Figure 18
Équation de Cobb-Douglas



1. Produit intérieur brut.
Source : Statistique Canada.

Figure 19
Équation de productivité du travail



1. Produit intérieur brut.
Source : Statistique Canada.

8. *Conclusion*

Peu importe la source de données utilisée, il est important d'examiner les données sous-jacentes afin de comprendre comment certaines caractéristiques des données, telles que les points aberrants et les points leviers, peuvent influer sur l'estimation et, par conséquent, sur l'inférence. Dans le cas de la base de données KLEMS (capital, travail, énergie, matières et services), cet examen est très important. Les données sous-jacentes sont bruitées et soumises aux changements méthodologiques, à l'erreur de mesure ainsi qu'aux chocs économiques agrégés et idiosyncratiques. Ces problèmes font en sorte qu'il est difficile de déterminer les « vraies » estimations paramétriques sous-jacentes en se servant de techniques d'estimation courantes telles que la méthode des moindres carrés ordinaires (MCO).

Le problème est amplifié lorsque les données sont utilisées pour effectuer des tests de racine unitaire parce que les tests couramment utilisés, comme le test de Dickey-Fuller amélioré (DFA) ou le test KPSS (Kwaitkowski, Phillips, Schmidt et Shin), ont une puissance peu élevée. De fait, ils ne parviennent pas à faire la distinction entre, d'une part, un processus de racine unitaire et, d'autre part, une réponse atypique ou un changement de tendance. Étant donné le bruit qu'il y a dans les données et le nombre d'observations atypiques présentes dans l'ensemble des variables et des industries de l'ensemble de données KLEMS, il est difficile d'appliquer les tests de racine unitaire les plus courants.

Lorsqu'on compare les estimations MCO et les estimations S, il semble que l'estimateur MCO sous-estime la croissance de *ptf* en raison des observations atypiques dans l'ensemble de données KLEMS. Les périodes où les observations atypiques sont présentes suggèrent que les chocs agrégés, telles les récessions, pourraient être à l'origine d'une proportion importante des

points aberrants et des points leviers que comporte l'ensemble de données. L'examen des périodes où la majorité des observations atypiques sont présentes et de la dispersion des différences logarithmiques de la production industrielle et du produit intérieur brut vient étayer cette hypothèse.

Les périodes où les observations atypiques sont présentes et l'incidence de celles-ci sur les estimations MCO suggèrent également que ce ne sont pas toutes les observations dans le temps qu'il convient d'utiliser pour estimer la productivité totale des facteurs (PTF). Cette dernière est souvent considérée comme la valeur ajoutée générée dans le processus de production lorsque les facteurs tangibles sont pris en considération. Cette interprétation convient bien aux périodes d'activité économique « normale ». Cependant, au cours des périodes de stress, comme le choc pétrolier du début des années 1970 et les récessions de 1980-1981 et de 1990, il est possible que l'interprétation ne corresponde pas aussi étroitement à l'évolution des facteurs et de la valeur ajoutée.

Notamment, des activités telles que le maintien d'effectifs en surnombre peuvent entraîner des changements dans les niveaux des facteurs qui ne sont pas équivalents aux changements dans la valeur ajoutée. En outre, la notion que l'économie réagit différemment selon qu'il s'agit de périodes d'expansion ou de récession est formulée explicitement dans les modèles de séries chronologiques avec changement de régime tels que le modèle markovien proposé par Hamilton (1989).

Par conséquent, lorsque les chercheurs estiment la PTF, ils seraient peut-être bien avisés de se demander si toutes les périodes fournissent une information d'égale valeur. La réponse dépendra dans une large mesure de la forme fonctionnelle présumée de la fonction de la PTF et de l'interprétation que les chercheurs choisissent de donner aux estimations de la PTF.

Bien que les chocs agrégés tels que le premier choc pétrolier, la récession de 1980-1981 et la récession de 1991 semblent expliquer une bonne partie des observations atypiques, ils ne les n'expliquent pas toutes. Les chocs idiosyncratiques, l'erreur de mesure et les changements méthodologiques sont probablement à l'origine des autres points aberrants et points leviers, mais il ne s'agit pour le moment que d'une supposition.

Quoi qu'il en soit, l'analyse montre qu'un estimateur robuste peut fournir des estimations paramétriques de la croissance de *ptf* à partir d'une base de données bruitée comportant différents types de points aberrants et de points leviers.

Annexe A Points leviers et points aberrants selon l'équation de Cobb-Douglas

Équation de Cobb-Douglas		
	Points aberrants	Points leviers
Cultures agricoles et élevage	...	1979, 1981, 1985, 1986
Foresterie et exploitation forestière	1975, 1976, 1983, 1987	1964, 1965, 1969, 1973, 1974, 1975, 1982, 1995, 1996
Pêche, chasse et piégeage	1968, 1994	1965, 1967, 1978, 1979, 1980, 1981, 1986, 1988, 1992
Activités de soutien à l'agriculture et à la foresterie	1987	1964, 1965, 1973, 1974, 1980, 1982, 1983, 1991, 1997, 1999, 2003
Extraction de pétrole et de gaz	1975	1966, 1968, 1980, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 2001
Extraction de charbon	1968, 1984, 1992, 2002, 2003	1962, 1966, 1967, 1968, 1970, 1971, 1980, 1982, 1983, 2000
Extraction de minerais métalliques	1964, 1973, 1984	1966, 1967, 1969, 1970, 1971, 1972
Extraction de minerais non métalliques	1982, 1984, 2003	1975, 1982, 1991, 1996, 1998, 2000
Activités de soutien à l'extraction minière et à l'extraction de pétrole et de gaz	1962, 1987, 1988	1992
Production, transport et distribution d'électricité	1982	1966, 1979, 1984, 1989, 1999
Distribution de gaz naturel et réseaux d'aqueduc et d'égout et autres	1962, 1997, 2002	1975, 1986, 1993, 1997, 1999, 2000, 2001, 2002, 2003
Construction	1978, 1980, 1991	1962, 1965, 1967, 1976, 1977, 1982, 1987, 1991
Fabrication d'aliments pour animaux	1973, 1974, 1982	1967, 1980, 1982, 1983, 1988, 1996, 1998, 2000, 2002, 2003
Fabrication de sucre et de confiseries	1965, 1974	1964, 1974, 1980, 1981, 1983, 1985, 1994, 1996, 2000, 2001, 2002
Mise en conserve de fruits et de légumes et fabrication de spécialités alimentaires	...	1977, 1992, 1996, 1998
Fabrication de produits laitiers	1971	1963, 1980, 1996, 2000
Fabrication de produits de viande	2000	1982, 1996, 1997, 1998, 1999, 2001, 2002, 2003
Préparation et conditionnement de poissons et de fruits de mer	1974, 1978	1962, 1966, 1967, 1978, 1988

Points leviers et points aberrants selon l'équation de Cobb-Douglas (suite)

Équation de Cobb-Douglas		
	Points aberrants	Points leviers
Fabrication de boissons gazeuses et de glace	1992	1967, 1970, 1997, 1998, 2001
Brasseries	...	1968, 1969, 1970, 1990, 1996, 1998, 2002, 2003
Vineries	1964, 1975, 1976	1994, 1998, 2001
Distilleries	1996, 1997	1969, 1970, 1978, 1979, 1980, 1992, 1996, 1997, 1999, 2001, 2002, 2003
Fabrication du tabac	1994, 1999	1974, 1975, 1986, 1987, 2000, 2001, 2002, 2003
Fabrication de textiles et de produits textiles	1981	1964, 1965, 1966, 1977, 1982, 1983, 1992, 2002, 2003
Fabrication de vêtements	1988, 1999, 2001	1977, 1982, 1983, 1986, 1990, 1991, 1992, 2000, 2001, 2002, 2003
Fabrication de produits en cuir et de produits analogues	2000	1972, 1981, 1982, 1992, 1993, 1999, 2001, 2002
Fabrication de produits en bois	1981, 1983	1969, 1970, 1974, 1975, 1982, 1990, 1991
Usines de pâte à papier, de papier et de carton	1975, 1983, 2001	1966, 1975, 1976, 1982, 1989, 1998
Fabrication de produits en papier transformé	1975, 1982, 1993	1975, 1977, 1982, 1985, 1988, 1989, 1994, 1996
Impression et activités connexes de soutien	1976, 1991	1983, 1989, 1990, 1991, 2003
Fabrication de produits du pétrole et du charbon	1962, 1973, 1982, 1983	1967, 1982, 1993, 2003
Fabrication de produits chimiques de base	1975, 1991	1975, 1976, 1983, 2003
Fabrication de résines, de caoutchouc synthétique et de fibres et de filaments artificiels et synthétiques	1981, 1982, 1983, 2001, 2003	1975, 1976, 1981, 1982, 1991, 1993, 1998, 1999, 2000, 2003
Fabrication de pesticides, d'engrais et d'autres produits chimiques agricoles	1974, 2001, 2003	1965, 1966, 1967, 1979, 1980, 1981, 1982, 1983, 1996, 1997, 1999
Fabrication de produits pharmaceutiques et de médicaments	1991, 2001	1966, 1979, 1982, 2001, 2003
Fabrication de produits divers chimiques	1981, 1982, 1997	1965, 1966, 1971, 1983, 1991, 1993, 1994, 1996, 1998, 2002, 2003

Points leviers et points aberrants selon l'équation de Cobb-Douglas (suite)

Équation de Cobb-Douglas		
	Points aberrants	Points leviers
Fabrication de produits en caoutchouc	1980, 1986, 1997	1970, 1971, 1981, 1982, 1989, 1990
Fabrication de ciment et de produits en béton	1967, 1997	1964, 1965, 1966, 1973, 1982, 1983, 1984, 1991, 1992, 1998
Fabrication de produits minéraux divers non métalliques	1980	1966, 1967, 1970, 1982, 1990, 1991, 1992, 1993
Première transformation des métaux	1979, 1984	1982, 1983, 1990, 1992, 1993, 1994, 2001, 2002, 2003
Fabrication de produits métalliques	1975, 1986, 1996	1965, 1966, 1982, 1983, 1990, 1991, 1992, 2000
Fabrication de machines	1984, 1993	1982, 1983, 1990, 1991, 1992, 2002
Fabrication de matériel informatique et périphérique	1971, 1981, 1984, 2001	1971, 1972, 1975, 1976, 1982, 1983, 1992, 2000, 2002
Fabrication de produits électroniques	1972, 1999, 2001, 2002	1966, 1982, 1999, 2001, 2002, 2003
Fabrication d'appareils ménagers	1975, 1987, 1988	1971, 1978, 1982, 1991
Fabrication de matériel et de composants électriques	2000	1966, 1982, 1983, 1991, 1998, 1999, 2000
Fabrication de véhicules automobiles	1967, 1980, 1988	1963, 1964, 1965, 1983, 1985, 1986, 1988
Fabrication de carrosseries et de remorques de véhicules automobiles	1964, 1982	1965, 1975, 1976, 1977, 1980, 1981, 1982, 1983, 1990, 1991, 2001
Fabrication de pièces pour véhicules automobiles	2001	1964, 1965, 1966, 1970, 1980, 1994
Fabrication de produits aérospatiaux et de leurs pièces	1983, 1984	1966, 1971, 1972, 1978, 1979, 1980, 1982, 2002, 2003
Fabrication de matériel ferroviaire roulant	1994, 2002, 2003	1963, 1975, 2001, 2003
Construction de navires et d'embarcations	2000, 2002	1962, 1964, 1971, 1979, 1981, 1982, 1991, 1999, 2001, 2002
Fabrication d'autres types de matériel de transport	1981, 1982, 2001, 2003	1965, 1967, 1968, 1969, 1970, 1982, 1994, 1995, 2001, 2003
Fabrication de meubles et de produits connexes	1974, 1981, 1982	1977, 1982, 1990, 1991, 1992

Points leviers et points aberrants selon l'équation de Cobb-Douglas (suite)

Équation de Cobb-Douglas		
	Points aberrants	Points leviers
Commerce de gros	1982	1963, 1982, 1983, 1992, 1994, 1997, 2001
Commerce de détail	1980, 1982, 1990, 1991	1982, 1991
Transport aérien	1991	1966, 1982, 1983, 1991
Transport ferroviaire	1982, 1984	1982, 1983, 1990, 1997, 1998, 1999
Transport par eau	1989	1981, 1982, 1983, 1986, 1996, 1998, 1999
Transport par camion	1974, 1979, 1980, 1991	1962, 1970, 1971, 1982, 1983, 1984, 1997, 1998
Transport en commun et transport terrestre de voyageurs	1979, 1981, 1984, 1991	1964, 1965, 1966, 1975, 1976, 1977, 1982
Transport par pipeline	1972, 1980	1966, 1969, 1978, 1980, 1981, 1998, 1999, 2000, 2001, 2003
Transport de tourisme et d'agrément et activités de soutien au transport	2000	1982, 1984, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003
Services postaux et messageries et services de messagers	1969, 1976, 1977, 1981, 1993	1963, 1965, 1967, 1968, 1980, 1981, 1992, 1994
Entreposage	1962, 1963	1996, 1998, 1999, 2001, 2002
Industries du film et de l'enregistrement sonore	1969, 1981, 1993	1967, 1968, 1982, 1991, 1992, 1993, 1997, 2001, 2002, 2003
Radiodiffusion et télécommunications	1975	1966, 1973, 1975, 1988, 2003
Édition et services d'information et de traitement des données	1991	1968, 1988, 1989, 1990, 1991, 1997, 1998, 1999, 2002
Autorités monétaires et intermédiation financière par le biais de dépôts	...	1981, 1987, 1991, 1992, 1993, 1998, 1999, 2002, 2003
Sociétés d'assurance	1996	1984, 1993, 1994, 2003
Bailleurs de biens immobiliers	1970, 1994, 1999	1977, 1986, 1987, 1992, 1993, 1997, 1998
Services de location et de location à bail et bailleurs de biens incorporels non financiers	1966	1964, 1982, 1991, 1997, 1999

Points leviers et points aberrants selon l'équation de Cobb-Douglas (fin)

Équation de Cobb-Douglas		
	Points aberrants	Points leviers
Autres finances, assurances et services immobiliers et gestion de sociétés et d'entreprises	1981, 1985, 1988, 1990	1972, 1976, 1977, 1978, 1984, 1988, 1993, 1994, 1997, 2002
Services de publicité et services connexes	1985, 1999	1966, 1976, 1992, 1994, 1995, 1999, 2001
Services d'architecture, d'ingénierie, juridique et de comptabilité	1974, 1992	1965, 1966, 1982, 1992, 1994
Autres services professionnels, scientifiques et techniques	1971, 1980	1965, 1966, 1971, 1982, 1983, 1992, 1998, 1999, 2002
Services administratifs et services de soutien	1968, 1988	1962, 1964, 1966, 1983, 1992, 1998, 1999, 2002
Services de gestion des déchets et d'assainissement	1965, 1967, 1974	1966, 1967, 1970, 1971, 1987, 1991
Services d'enseignement, sauf les universités	1967, 1974, 1996, 1997	1962, 1963, 1975, 1976, 1984, 1993, 1994, 1996, 1997, 1998, 1999
Soins de santé et assistance sociale, sauf les hôpitaux	1975	1963, 1964, 1965, 1966, 1971, 1976, 1999, 2001, 2002, 2003
Arts, spectacles et loisirs	1974, 1987	1963, 1965, 1966, 1967, 1968, 1972, 1973, 1975, 1980, 1981, 1989
Hébergement et services de restauration	1982, 1991	1963, 1965, 1966, 1968, 1972, 1977, 1982, 1991, 2003
Réparation et entretien	1969, 1974, 1984	1966, 1982, 1991, 1998, 1999, 2000
Fondations, groupes de citoyens et organisations professionnelles et similaires	1962, 1963, 1965, 1967, 1974	1964, 1965, 1966, 1971, 1974, 1981, 1992, 1993, 1995, 1996, 1997
Services personnels et services de blanchissage et ménages privés	1981, 1984, 1985	1966, 1980, 1991, 1994, 1995, 1996, 1998
Secteur non commercial	1968	1962, 1965, 1966, 1967, 1996, 1997, 1998, 2002

... n'ayant pas lieu de figurer

Annexe B Points leviers et points aberrants selon l'équation de productivité du travail

Équation de productivité du travail		
	Points aberrants	Points leviers
Cultures agricoles et élevage	1967	1985
Foresterie et exploitation forestière	1974, 1975, 1976, 1983, 1987	1975, 1978, 1995, 1996
Pêche, chasse et piégeage	1968, 1994	1965, 1980, 1986, 1988
Activités de soutien à l'agriculture et à la foresterie	1987, 1997	1974, 1997
Extraction de pétrole et de gaz	1975, 1980, 1981	1966, 1968, 2001
Extraction de charbon	1968, 1984	1962, 1967, 1968, 1983
Extraction de minerais métalliques	1971, 1973, 1984	1969, 1982
Extraction de minerais non métalliques	1982, 1984, 2003	1982, 1997
Activités de soutien à l'extraction minière et à l'extraction de pétrole et de gaz	1962, 1986, 1987, 1988	1994, 2000
Production, transport et distribution d'électricité	1975, 1981, 1982, 1989	1966, 1989, 2000
Distribution de gaz naturel et réseaux d'aqueduc et d'égout et autres	1962, 1978	1975, 1997, 2002
Construction	1978, 1980, 1991	1962, 1967, 1976, 1982, 1987, 1991
Fabrication d'aliments pour animaux	1973, 1974, 1982	1980, 1983, 1988, 1996, 2000
Fabrication de sucre et de confiseries	1965, 1974, 1980	1964, 1974, 1981, 1985, 1994, 2000, 2001, 2002
Mise en conserve de fruits et de légumes et fabrication de spécialités alimentaires	...	1992, 1996, 1998
Fabrication de produits laitiers	1971, 1980, 2002	1980
Fabrication de produits de viande	1962, 1989, 2000	1997, 1998, 1999, 2001, 2003
Préparation et conditionnement de poissons et de fruits de mer	1974	1976, 1977, 2000
Fabrication d'aliments divers	1976, 1991, 1999, 2001	1964, 1970

Points leviers et points aberrants selon l'équation de productivité du travail (suite)

Équation de productivité du travail		
	Points aberrants	Points leviers
Fabrication de boissons gazeuses et de glace	1992, 2002	1970, 1997, 1998, 1999, 2001
Brasseries	...	1969, 1990, 1998, 2002, 2003
Vineries	1964, 1975, 1976	2001
Distilleries	1997	1970, 1979, 1980, 1999, 2001, 2002, 2003
Fabrication du tabac	1974, 1975, 1994	1974, 1975, 1986, 2001, 2003
Fabrication de textiles et de produits textiles	1981, 1982	1983, 2002
Fabrication de vêtements	1988, 1999, 2001	1983, 1993, 1997, 2000, 2001
Fabrication de produits en cuir et de produits analogues	2000	1981, 1982, 1991, 1993, 2002
Fabrication de produits en bois	1983	1970, 1974
Usines de pâte à papier, de papier et de carton	1975, 1983, 2001	1975, 1982, 1988, 1989
Fabrication de produits en papier transformé	1993	1975, 1982, 1985, 1988, 1989, 1996
Impression et activités connexes de soutien	1976, 1991	1983, 1989, 1990, 1991, 2003
Fabrication de produits du pétrole et du charbon	1962, 1982, 1983	1982, 1993
Fabrication de produits chimiques de base	1975, 1991	1983, 1995, 2003
Fabrication de résines, de caoutchouc synthétique et de fibres et de filaments artificiels et synthétiques	1981, 1983, 2001, 2003	1975, 1976, 1993, 1999, 2000, 2003
Fabrication de pesticides, d'engrais et d'autres produits chimiques agricoles	1974, 2003	1966, 1979, 1980, 1983, 1997
Fabrication de produits pharmaceutiques et de médicaments	1991, 1996, 2001, 2003	1966, 1979, 1982, 1994, 2001
Fabrication de produits divers chimiques	1981, 1982, 1997	1983, 1991, 1998, 2003
Fabrication de produits en plastique	1980, 1981	1982

Points leviers et points aberrants selon l'équation de productivité du travail (suite)

Équation de productivité du travail		
	Points aberrants	Points leviers
Fabrication de produits en caoutchouc	1980, 1984, 1986, 1993, 1997	1970, 1971, 1981, 1982, 1989, 1990
Fabrication de ciment et de produits en béton	1967, 1997	1971, 1986, 1987
Fabrication de produits minéraux divers non métalliques	1980	1966, 1967, 1970, 1982, 1990
Première transformation des métaux	1979, 1983, 1984	1975, 1982, 1990, 2002
Fabrication de produits métalliques	1975	1962, 1982, 1992, 1993, 2000
Fabrication de machines	1984, 1993, 2002	1982, 1990, 1991
Fabrication de matériel informatique et périphérique	1971, 1981, 1984, 2001	1971, 1972, 1975, 1992, 2000, 2002
Fabrication de produits électroniques	1972, 1999, 2001, 2002	1999, 2001, 2002, 1982, 2003
Fabrication d'appareils ménagers	1975, 1987, 1988	1987, 1971, 1978, 1991
Fabrication de matériel et de composants électriques	2000	2000, 1982, 1983, 1991, 1998
Fabrication de véhicules automobiles	1967, 1980, 1988	1971, 1986
Fabrication de carrosseries et de remorques de véhicules automobiles	1964, 1982	1982, 1965, 1991
Fabrication de pièces pour véhicules automobiles	2001	1970, 1980
Fabrication de produits aérospatiaux et de leurs pièces	1983, 1984	1972
Fabrication de matériel ferroviaire roulant	1994, 2002, 2003	1963, 2002, 2001, 2003
Construction de navires et d'embarcations	1997	1962, 1991, 2001, 2002
Fabrication d'autres types de matériel de transport	1981, 1982, 2001, 2003	1982, 2001, 2003
Fabrication de meubles et de produits connexes	1974, 1981, 1982	...
Activités diverses de fabrication	1984, 1986, 1997	1972, 1995, 2000

Points leviers et points aberrants selon l'équation de productivité du travail (suite)

Équation de productivité du travail		
	Points aberrants	Points leviers
Commerce de gros	1982	1994
Commerce de détail	1979, 1980, 1986, 1990, 1991	1991, 1976, 1987, 1993, 2000
Transport aérien	1991	1991
Transport ferroviaire	1982	1982, 1983, 1997, 1999
Transport par eau	1989	1982, 1983, 1993, 1994, 1996, 1999
Transport par camion	1974, 1979	1962, 1970, 1971, 1983, 1984
Transport en commun et transport terrestre de voyageurs	1979, 1981, 1984	1964, 1965, 1975, 1976, 1982
Transport par pipeline	1980, 1981, 1982	1980, 1966, 1978, 1998, 2001, 2003
Transport de tourisme et d'agrément et activités de soutien au transport	2000, 2002	1982, 1984, 1998
Services postaux et messageries et services de messagers	1969, 1976, 1977, 1981	1965, 1968, 1994
Entreposage	1998	1998, 1996, 2001
Industries du film et de l'enregistrement sonore	1969, 1993	1967, 1968,
Radiodiffusion et télécommunications	1975	1975, 2003
Édition et services d'information et de traitement des données	1991	1989, 1990, 1991, 1998, 2002
Autorités monétaires et intermédiation financière par le biais de dépôts	...	1987, 1992, 1998, 1999
Sociétés d'assurance	1996	1993, 1999
Bailleurs de biens immobiliers	1994, 1999	1977, 1986, 1993, 1998
Services de location et de location à bail et bailleurs de biens incorporels non financiers	1966, 2001	1962, 1963, 1964, 1977, 1997
Autres finances, assurances et services immobiliers et gestion de sociétés et d'entreprises	1981, 1985, 1988	1976, 1978, 1984, 1994, 1997, 2002

Points leviers et points aberrants selon l'équation de productivité du travail (fin)

Équation de productivité du travail		
	Points aberrants	Points leviers
Services de publicité et services connexes	1976, 1977	1966, 1994, 1995, 1999, 2001
Services d'architecture, d'ingénierie, juridique et de comptabilité	...	1965, 1966, 1982, 1994
Autres services professionnels, scientifiques et techniques	1971, 1980	1965, 1966, 1992, 1998
Services administratifs et services de soutien	1968, 1988	1962, 1966, 1998, 1999, 2002
Services de gestion des déchets et d'assainissement	1965, 1974	1966, 1970, 1971
Services d'enseignement, sauf les universités	1962, 1974, 1996, 1998	1962, 1996, 1998, 1975, 1976, 1984, 1999
Soins de santé et assistance sociale, sauf les hôpitaux	1973	1976, 1999, 2001, 2002, 2003
Arts, spectacles et loisirs	1974, 1987	1965, 1966, 1967, 1968, 1972, 1975, 1980
Hébergement et services de restauration	1972, 1977	1972, 1965, 1966
Réparation et entretien	1969, 1974, 1984, 1991	1966, 1998, 1999, 2000, 2001
Fondations, groupes de citoyens et organisations professionnelles et similaires	1962, 1963, 1965, 1967, 1974	1965, 1964, 1981, 1993, 1996, 1997
Services personnels et services de blanchissage et ménages privés	1997	1966, 1995, 1996, 1998
Secteur non commercial	1968	1967

... n'ayant pas lieu de figurer

Bibliographie

Baldwin, John R., René Durand et Judy Hosein. 1999. « Restructuration et croissance de la productivité dans le secteur des entreprises au Canada ». Dans *Croissance de la productivité au Canada.* 27–40. N° 15-204-XPF1999000 au catalogue. Ottawa : Statistique Canada.

Baldwin, John R., et Wulong Gu. 2007. « La productivité multifactorielle au Canada : une évaluation de diverses méthodes d'estimation des services de capital ». Dans *La revue canadienne de productivité.* N° 15-206-XIF2007009 au catalogue. Ottawa : Statistique Canada.

Balke, Nathan S., et Thomas B. Fomby. 1994. « Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series. » *Journal of Applied Econometrics.* 9, 2 : 181–200.

Beckstead, Desmond, et Jean-Pierre Maynard. 2001. « Annexe 2—Classification industrielle des measures de productivité et variables connexes ». Dans *Croissance de la productivité au Canada.* 187–194. N° 15-204-XPF au catalogue. Ottawa : Statistique Canada.

Chang, Ih, George C. Tiao et Chung Chen. 1988. « Estimation of time series parameters in the presence of outliers ». *Technometrics.* 30, 2 : 193–204.

Chen, Chung, et Lon-Mu Liu. 1993. « Joint estimation of model parameters and outlier effects in time series ». *Journal of the American Statistical Association.* 88, 421 : 284–297.

Chen, Colin. 2002. « Robust Regression and Outlier Detection with the ROBUSTREG Procedure ». SAS Conference Proceedings: SUGI 27. Document n° 265-27. SAS Users Group International.

Farley, Edward V., et James M. Murphy. 1997. « Time series outlier analysis: Evidence for management and environmental influences on sockeye salmon catches in Alaska and northern British Columbia ». *Alaska Fishery Research Bulletin.* 4, 1 : 36–53.

Fox, A.J. 1972. « Outliers in time series ». *Journal of the Royal Statistical Society. Series B (Methodological).* 34, 3 : 350–363.

Franses, Philip Hans. 1998. « Time Series Models for Business and Economic Forecasting ». New York : Cambridge University Press.

Fung, Wing-Kam. 1999. « Outlier diagnostics in several multivariate samples ». *The Statistician.* 48, 1 : 73–84.

Galeano, Pedro, Daniel Peña et Ruey S. Tsay. 2006. « Outlier detection in multivariate time series by project pursuit ». *Journal of the American Statistical Association.* 101, 474 : 654–669.

Gellatly, Guy, Marc Tanguay et Beiling Yan. 2002. « Une méthode alternative d'estimation de la dépréciation économique : nouveaux résultats obtenus au moyen d'un modèle de survie ». Dans *Croissance de la productivité au Canada - 2002*. 25–65. John R. Baldwin et Tarek M. Harchaoui (éditeurs). N° 15-204-XIF2001000 au catalogue. Ottawa : Statistique Canada.

Granger, C.W.J., et P. Newbold. 1974. « Spurious regressions in economics ». *Journal of Econometrics*. 2, 2 : 111–120.

Gu, Wulong, Mustapha Kaci, Jean-Pierre Maynard et Mary-Anne Sillamaa. 2002. « Changement de la composition de la population active canadienne et son influence sur la croissance de la productivité ». Dans *Croissance de la productivité au Canada*. 71–98. John R. Baldwin et Tarek M. Harchaoui (éditeurs). N° 15-204-XIF2001000 au catalogue. Ottawa : Statistique Canada.

Hadi, Ali S. 1992. « Identifying multiple outliers in multivariate data ». *Journal of the Royal Statistical Society. Series B (Methodological)*. 54, 3 : 761–771.

Hadi, Ali S. 1994. « A modification of a method for the detection of outliers in multivariate samples ». *Journal of the Royal Statistical Society. Series B (Methodological)*. 56, 2 : 393–396.

Hamilton, James D. 1989. « A new approach to the economic analysis on nonstationary time series and the business cycle ». *Econometrics*. 57, 2 : 357–384.

Hadri, Kaddour. 2000. « Testing for stationarity in heterogeneous panel data ». *Econometrics Journal*. 3, 2 : 148–161.

Harchaoui, Tarek M., et Faouzi Tarkhani. 2003. « Une révision complète de la méthode d'estimation de l'intrant capital pour le programme de la productivité multifactorielle de Statistique Canada ». Dans *Croissance de la productivité au Canada - 2002*. 107–161. John R. Baldwin et Tarek M. Harchaoui (éditeurs). N° 15-204-XIF20001000 au catalogue. Ottawa : Statistique Canada.

Im, Kyung So, M. Hashem Pesaran et Yongcheol Shin. 1995. « Testing for Unit Roots in Heterogeneous Panels ». DAE Working Papers Amalgamated Series n° 9526. Cambridge, R.U.: University of Cambridge. (Aussi dans *Journal of Econometrics*. 2003. 115, 1 : 53–74.)

Jorgenson, Dale W., Frank M. Gollop et Barbara M. Fraumeni. 1987. « Productivity and U.S. Economic Growth ». Harvard Economic Studies. Cambridge, Mass.: Harvard University Press.

Jorgenson, Dale W., et Kevin J. Stiroh. 2000. « U.S. economic growth at the industry level ». *The American Economic Review*. 90, 2 : 161–167. Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association.

Kwaitkowski, Denis, Peter C.B. Phillips, Peter Schmidt et Yongcheol Shin. 1992. « Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? » *Journal of Econometrics*. 54, 1–3 : 159–178.

Ljung, Greta M. 1993. « On outlier detection in time series ». *Journal of the Royal Statistical Society. Series B. (Methodological).* 55, 2 : 559–567.

Maddala, G.S., et In-Moo Kim. 1998. « Unit Roots, Cointegration, and Structural Change ». New York : Cambridge University Press.

Maddala, G.S., et Shaowen Wu. 1999. « A comparative study of unit root tests with panel data and a new simple test ». *Oxford Bulletin of Economics and Statistics.* Special Issue. 16, 4 : 631–652.

Nelson, C.R., et C.I. Plosser. 1982. « Trends and random walks in macroeconomic time series: some evidence and implications ». *Journal of Monetary Economics.* 10 : 139–162.

Peña, Daniel. 1990. « Influential observations in time series ». *Journal of Business & Economic Statistics.* 8, 2 : 235–241.

Perron, Pierre. 1989. « The great crash, the oil price shock, and the unit root hypothesis ». *Econometrica.* 57, 6 : 1361–1401.

Rousseeuw, Peter J. 1984. « Least median of squares regression ». *Journal of the American Statistical Association.* 79, 388 : 871–880.

---------. 1985. « Multivariate estimation with a high breakdown point ». Dans *Mathematical Statistics and Applications. Volume B.* 283–297. W. Grossman, G. Pflung, I. Vincze et W. Wertz (réviseurs). Dordrecht, Hollande : Reidel Publishing Company.

Rousseeuw, Peter J., et Annick M. Leroy. 1987. « Robust Regression and Outlier Detection ». New York : Wiley.

Rousseeuw, Peter J., et Bert C. van Zomeren. 1990. « Unmasking multivariate outliers and leverage points ». *Journal of the American Statistical Association.* 85, 411 : 633–639.

Rousseeuw, Peter J., et Katrien van Driessen. 1999. « A fast algorithm for the minimum covariance determinant estimator ». *Technometrics.* 41, 3 : 212–223.

Rousseeuw, Peter J., et Victor J. Yohai. 1984. « Robust regression by means of S-estimators ». Dans *Robust and Nonlinear Time Series Analysis.* Notes de cours dans Statistics n° 26. New York : Springer.

Ruppert, David. 1992. « Computing S estimators for regression and multivariate location/dispersion ». *Journal of Computational and Graphical Statistics.* 1, 3 : 253–270.

Sapra, Sunil K. 2003. « High-breakdown point estimation of some regression models ». *Applied Economics Letters.* 10, 14 : 875–878.

Schreyer, Paul, W. Erwin Diewert et Anne Harrison. 2005. « Cost of Capital Services and the National Accounts ». Étude technique pour la réunion de la Experts Advisory Group on National Accounts(SNA/M1.05/04).
http://unstats.un.org/unsd/nationalaccount/AEG/papers/m3CapitalServices.pdf (consulté le 17 juillet 2007.)

Tsay, Ruey S. 1986. « Time series model specification in the presence of outliers ». *Journal of the American Statistical Association*. 81, 393 : 132–141.

Tsay, Ruey S., Daniel Peña et Alan E. Pankratz. 2000. « Outliers in multivariate time series ». *Biometrika*. 87, 4 : 789–804.

Yule, G. Udny. 1926. « Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time-series ». *Journal of the Royal Statistical Society*. 89, 1 : 1–64.